

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:25:47 ; Search time 23 Seconds
(without alignments)
20.201 Million cell updates/sec

Title: US-09-756-899A-1
Perfect score: 63
Sequence: 1 AHWSGHCCL 9

Scoring table: BLOSUM62
Gapop 13.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 89883

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued Patents AA:*
- 1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
 - 2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
 - 3: /cgn2_6/prodata/2/iaa/6A COMB.pep:*
 - 4: /cgn2_6/prodata/2/iaa/6B COMB.pep:*
 - 5: /cgn2_6/prodata/2/iaa/PCUTUS COMB.pep:*
 - 6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	55.6	8	2	US-08-557-122A-24
2	35	55.6	8	4	US-09-262-666-24
3	33	52.4	6	1	US-08-068-395A-4
4	33	52.4	6	1	US-08-464-365-4
5	33	52.4	6	3	US-08-627-907A-5
6	33	52.4	6	3	US-07-301-713A-1
7	33	52.4	7	2	US-08-323-536A-5
8	32	50.8	7	2	US-08-645-193B-65
9	28	44.4	8	3	US-08-657-339A-12
10	28	44.4	8	4	US-08-457-694A-12
11	26	41.3	4	1	US-07-609-716-10
12	26	41.3	4	1	US-08-477-509B-107
13	26	41.3	4	3	US-08-482-085B-107
14	26	41.3	4	3	US-08-475-411A-10
15	26	41.3	4	3	US-08-478-029A-10
16	26	41.3	4	4	US-09-444-791A-107
17	26	41.3	5	2	US-08-244-496-83
18	26	41.3	7	2	US-08-923-536A-6
19	26	41.3	9	3	US-09-258-754-368
20	26	41.3	9	3	US-09-042-107-368
21	26	41.3	9	4	US-09-722-2500-368
22	25	39.7	6	1	US-08-487-006-59
23	25	39.7	6	2	US-08-488-659A-59
24	25	39.7	6	4	US-09-058-740-6
25	25	39.7	8	2	US-08-520-535-12
26	25	39.7	8	2	US-09-079-432-12
27	25	39.7	9	3	US-09-518-046-63

Sequence 74, Appl
Sequence 76, Appl
Sequence 78, Appl
Patent No. 5217955
Sequence 4, Appl
Sequence 5, Appl
Sequence 3, Appl
Sequence 17, Appl
Sequence 20, Appl
Sequence 17, Appl
Sequence 20, Appl
Sequence 3, Appl
Sequence 17, Appl
Sequence 20, Appl
Sequence 27, Appl
Sequence 8, Appl

28 25 39.7 9 4 US-09-270-542-74
29 25 39.7 9 4 US-09-270-542-76
30 25 39.7 9 4 US-09-270-542-78
31 24.5 38.9 9 6 5217955-26
32 24 38.1 4 2 US-07-733-095B-4
33 24 38.1 5 1 US-07-826-328A-4
34 24 38.1 5 1 US-07-826-328A-5
35 24 38.1 5 1 US-07-973-235A-3
36 24 38.1 5 1 US-08-350-884-17
37 24 38.1 5 1 US-08-350-884-20
38 24 38.1 5 1 US-08-709-173-17
39 24 38.1 5 1 US-08-709-173-20
40 24 38.1 5 2 US-08-462-720-3
41 24 38.1 5 2 US-08-709-177-17
42 24 38.1 5 2 US-08-709-177-20
43 24 38.1 5 4 US-08-871-561-27
44 24 38.1 5 4 US-09-321-932B-27
45 24 38.1 6 1 US-07-836-928A-8

ALIGNMENTS

RESULT 1
US-08-557-122A-24
; Sequence 24, Application US/08557122A
; Patent No. 5879664
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Mailand
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5879664 No. 5879664disk of No. 5879664th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,122A
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-557-122A-24

Query Match 55.6%; Score 35; DB 2; Length 8;
Best Local Similarity 71.4%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AHWSGHC 7
Db 1 AFWCGHC 7

RESULT 2

US-09-262-666-24
; Sequence 24, Application US/09262666
; Patent No. 6346244
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Mailand
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 63462440 No. 6346244disk of No. 6346244th America, Inc.
; STREET: 405 Lexington Avenue, 54th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/262,666
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,122
; FILING DATE: 11-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-262-666-24

Query Match 55.6%; Score 35; DB 4; Length 8;
Best Local Similarity 71.4%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AHWGHC 7
DB 1 APWCGHC 7

RESULT 3
US-08-068-395A-4
; Sequence 4, Application US/08068395A
; Patent No. 5496719
; GENERAL INFORMATION:
; APPLICANT: YAMADA, YUKIO
; APPLICANT: ASAMI, OSAMU
; APPLICANT: SUGIYAMA, HIDEHIKO
; APPLICANT: IDEKOB, CHIE
; APPLICANT: HOSHINO, FUMIHIKO
; APPLICANT: HIRAI, MASANA
; APPLICANT: KAJINO, TSUTOMU
; APPLICANT: IMAEDA, TAKAO
; APPLICANT: SARAI, KIYOKO
; TITLE OF INVENTION: POLYPEPTIDE POSSESSING PROTEIN DISULFIDE
; TITLE OF INVENTION: ISOMERASE ACTIVITY GENE ENCODING THE SAME
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/068,395A
FILING DATE: 19930527
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-135254
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-44013
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-44014
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5496719man P.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 68-228-0
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-068-395A-4

Query Match 52.4%; Score 33; DB 1; Length 6;
Best Local Similarity 80.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WSGHC 7
DB 1 WCGHC 5

RESULT 4
US-08-464-365-4
; Sequence 4, Application US/08464365
; Patent No. 5700659
; GENERAL INFORMATION:
; APPLICANT: YAMADA, YUKIO
; APPLICANT: ASAMI, OSAMU
; APPLICANT: SUGIYAMA, HIDEHIKO
; APPLICANT: IDEKOB, CHIE
; APPLICANT: HOSHINO, FUMIHIKO
; APPLICANT: HIRAI, MASANA
; APPLICANT: KAJINO, TSUTOMU
; APPLICANT: IMAEDA, TAKAO
; APPLICANT: SARAI, KIYOKO
; TITLE OF INVENTION: POLYPEPTIDE POSSESSING PROTEIN DISULFIDE
; TITLE OF INVENTION: ISOMERASE ACTIVITY GENE ENCODING THE SAME
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,365
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-135254
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-44013
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-44014
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: OGILO, NO. 5700859man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 68-228-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-464-365-4

Query Match 52.4%; Score 33; DB 1; Length 6;
Best Local Similarity 80.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WSGHC 7
Db 1 WCGHC 5

RESULT 5
US-08-627-907A-5
Sequence 5, Application US/08627907A
Patent No. 6060302
GENERAL INFORMATION:
APPLICANT: HIRANO, Naoto
APPLICANT: HIRAI, Hisamaru
TITLE OF INVENTION: HUMAN PHOSPHOLIPASE C-ALPHA AND DNA
TITLE OF INVENTION: SEQUENCE ENCODING THE SAME
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSES: c/o FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/627,907A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-238402
FILING DATE: 24-SEP-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP94/01572
FILING DATE: 22-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: HALEY Jr., James P.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: SHGN-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-627-907A-5

Query Match 52.4%; Score 33; DB 3; Length 6;
Best Local Similarity 80.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WSGHC 7
Db 1 WCGHC 5

RESULT 6
US-07-901-713A-1
Sequence 1, Application US/07901713A
Patent No. 6291205
GENERAL INFORMATION:
APPLICANT: Tuite, Michael F.
APPLICANT: Freedman, Robert B.
APPLICANT: Markus, Henry Z.
APPLICANT: Schultz, Loren D.
APPLICANT: Montgomery, Donna L.
APPLICANT: Ellis, Ronald W.
TITLE OF INVENTION: METHOD FOR INCREASING PRODUCTION OF
TITLE OF INVENTION: DISULFIDE BONDED RECOMBINANT PROTEINS BY SACCAROMYCES
TITLE OF INVENTION: CERREVISIAE
FILE REFERENCE: 18469
CURRENT APPLICATION NUMBER: US/07/901,713A
CURRENT FILING DATE: 1992-06-12
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 6
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: vertebrate PDI active site
US-07-901-713A-1

Query Match 52.4%; Score 33; DB 3; Length 6;
Best Local Similarity 80.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WSGHC 7
Db 1 WCGHC 5

RESULT 7
US-08-923-536A-5
Sequence 5, Application US/08923536A
Patent No. 5965426
GENERAL INFORMATION:
APPLICANT: SAKAI, Yasuyoshi
APPLICANT: KATO, No. 5965426
APPLICANT: SHIBANO, YUJI
TITLE OF INVENTION: PROTEIN DISULFIDE ISOMERASE GENE DERIVED
TITLE OF INVENTION: FROM STRAIN OF METHYLOTROPIC YEAST

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWICKER & MATHIS, L.L.P.
STREET: 1737 King Street, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22314-2756
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,536A
FILING DATE: 04-SEP-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-234287
FILING DATE: 04-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Neuth, Donna M.
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 001560-317
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-923-536A-5

Query Match 52.4%; Score 33; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 WSGHC 7
Db 2 WGHCC 6

RESULT 8
US-08-645-193B-65
Sequence 65, Application US/08645193B
Patent No. 5962253
GENERAL INFORMATION:
APPLICANT: Kupke, Thomas
APPLICANT: Gotz, Friedrich
APPLICANT: Kempter, Christoph
APPLICANT: Jung, Gunther
TITLE OF INVENTION: Oxidative Decarboxylation of Peptides
TITLE OF INVENTION: Catalyzed by Flavoprotein EpiD
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/645,193B
FILING DATE: 13-MAY-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652.1540000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-645-193B-65

Query Match 50.8%; Score 32; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 WSGHCC 8
Db 2 WNSYCC 7

RESULT 9
US-08-657-339A-12
Sequence 12, Application US/08657339A
Patent No. 6265551
GENERAL INFORMATION:
APPLICANT: Duke-Cohan, Jonathan S.
APPLICANT: Morimoto, Chikao
APPLICANT: Schlossman, Stuart F.
TITLE OF INVENTION: NOVEL FORM OF DIPEPTIDYLPEPTIDASE IV (CD26) FOUND IN
TITLE OF INVENTION: HUMAN SERUM, ANTIBODIES THERETO, AND USES THEREFOR
FILE REFERENCE: 00530-091001
CURRENT APPLICATION NUMBER: US/08/657,339A
CURRENT FILING DATE: 1996-06-03
PRIOR APPLICATION NUMBER: US 08/457,694
PRIOR FILING DATE: 1995-06-01
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 12
LENGTH: 8
TYPE: PRT
ORGANISM: Homo sapiens
US-08-657-339A-12

Query Match 44.4%; Score 28; DB 3; Length 8;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 SGHCCL 9
Db 1 SGHICI 6

RESULT 10
US-08-457-694A-12
Sequence 12, Application US/08457694A
Patent No. 6325989
GENERAL INFORMATION:
APPLICANT: Duke-Cohan, Jonathan S.
APPLICANT: Morimoto, Chikao
APPLICANT: Schlossman, Stuart F.
TITLE OF INVENTION: NOVEL FORM OF DIPEPTIDYLPEPTIDASE IV FOUND IN
TITLE OF INVENTION: HUMAN SERUM
FILE REFERENCE: 00530-090001
CURRENT APPLICATION NUMBER: US/08/457,694A
CURRENT FILING DATE: 1995-06-01
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 12

```
;
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-457-694A-12

Query Match          44.4%; Score 28; DB 4; Length 8;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 4 SGHCL 9
Db 1 SGHIC 6

RESULT 11
US-07-609-716-10
; Sequence 10, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-609-716-10

Query Match          41.3%; Score 26; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HCC 8
Db 1 HCC 3

RESULT 12
US-08-477-509B-107
; Sequence 107, Application US/08477509B
; Patent No. 5770697
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W

; APPLICANT: Dorman, Mary A
; TITLE OF INVENTION: No. 5770697el Peptides Comprising Repetitive
; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,509B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,155
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,049
; FILING DATE: 22-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-7/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-477-509B-107

Query Match          41.3%; Score 26; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HCC 8
Db 1 HCC 3

RESULT 13
US-08-492-085B-107
; Sequence 107, Application US/08482085B
; Patent No. 6018030
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Richardson, Charles
; APPLICANT: Chambers, James
; APPLICANT: Causey, Stuart
; APPLICANT: Pollock, Thomas J.
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W
; TITLE OF INVENTION: No. 6018030el Peptides Comprising Repetitive
; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
 STREET: Four Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: US
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/482.085B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 06/927,258
 FILING DATE: 04-NOV-1986
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/114,618
 FILING DATE: 29-OCT-1987
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/053,049
 FILING DATE: 22-APR-1993
 APPLICATION NUMBER: US 08/175,155
 FILING DATE: 29-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Trecartin, Richard F.
 REGISTRATION NUMBER: 31,801
 REFERENCE/DOCKET NUMBER: A-55186-6/RFT/MTK
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-781-1989
 TELEFAX: 415-398-3249
 INFORMATION FOR SEQ ID NO: 107:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-482-085B-107

Query Match 41.3%; Score 26; DB 3; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 6 HCC 8
 Db 1 HCC 3

RESULT 14
 US-08-475-411A-10
 Sequence 10, Application US/08475411A
 Patent No. 6140072
 GENERAL INFORMATION:
 APPLICANT: Ferrari, Franco A.
 APPLICANT: Cappello, Joseph
 TITLE OF INVENTION: Functional Recombinantly Prepared
 TITLE OF INVENTION: Synthetic Protein Polymer
 NUMBER OF SEQUENCES: 119
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
 STREET: Four Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: CA
 COUNTRY: US
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/475,411A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/609,716
 FILING DATE: 06-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/269,429
 FILING DATE: 09-NOV-1988
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/114,618
 FILING DATE: 29-OCT-1987
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 06/927,258
 FILING DATE: 04-NOV-1986
 ATTORNEY/AGENT INFORMATION:
 NAME: Trecartin, Richard F.
 REGISTRATION NUMBER: 31,801
 REFERENCE/DOCKET NUMBER: A-55186-9/RFT/MTK
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-781-1989
 TELEFAX: 415-398-3249
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-475-411A-10

Query Match 41.3%; Score 26; DB 3; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 HCC 8
 Db 1 HCC 3

RESULT 15
 US-08-478-029A-10
 Sequence 10, Application US/08478029A
 Patent No. 6184348
 GENERAL INFORMATION:
 APPLICANT: Ferrari, Franco A.
 APPLICANT: Cappello, Joseph
 TITLE OF INVENTION: Functional Recombinantly Prepared
 TITLE OF INVENTION: Synthetic Protein Polymer
 NUMBER OF SEQUENCES: 119
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
 STREET: Four Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: CA
 COUNTRY: US
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/478,029A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/609,716
 FILING DATE: 06-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/269,429

```

; FILING DATE: 09-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard P.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-8/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-478-029A-10

Query Match 41.3%; Score 26; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HCC 8
Db 1 HCC 3

Search completed: June 1, 2004, 10:29:22
Job time : 24 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:25:46 ; Search time 21 Seconds
(without alignments)
41.225 Million cell updates/sec

Title: US-09-756-899A-1

Perfect score: 63

Sequence: 1 AHWGHCCCL 9

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 790

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:**

1: piri:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	25	39.7	8	2 D61512	variant surface gl
2	23	36.5	9	2 A60522	sperm-activating p
3	21	33.3	8	2 C61512	variant surface gl
4	18	28.6	6	2 PT0652	T-cell receptor be
5	18	28.6	6	4 I79564	hypothetical TGL3
6	18	28.6	8	2 FC1002	leucine-tRNA ligas
7	18	28.6	9	2 QDRB	delta sleep-induci
8	18	28.6	9	2 A43848	cell surface adhes
9	18	28.6	9	2 PT0562	T-cell receptor be
10	18	28.6	9	2 PH0462	T-cell receptor be
11	17	27.0	4	2 B53294	T-cell receptor be
12	17	27.0	7	2 S08606	hypothetical prote
13	17	27.0	7	2 PT0642	T-cell receptor be
14	17	27.0	7	2 PT0758	T-cell receptor be
15	17	27.0	7	2 PT0786	T-cell receptor be
16	17	27.0	8	2 A25836	L-serine ammonia-1
17	17	27.0	9	2 A12872	transaldolase (EC
18	17	27.0	9	2 A11497	transaldolase (EC
19	17	27.0	9	2 PT0634	T-cell receptor be
20	16	25.4	5	2 PT0281	Ig heavy chain CRD
21	15.5	24.6	9	2 S07205	litorin 2-Glu - Au
22	15.5	24.6	9	2 S07204	litorin I - Austr
23	15.5	24.6	9	2 S07241	litorin - Rohde's
24	15	23.8	4	2 PT0661	T-cell receptor be
25	15	23.8	5	2 B45525	actin I - malaria
26	15	23.8	9	2 PT0231	Ig heavy chain CRD
27	15	23.8	9	2 A37027	macrophage chemot
28	14	22.2	3	3 GKHU	growth-modulating
29	14	22.2	4	2 FL0140	carbon-monoxide de

ALIGNMENTS

RESULT 1

D61512

variant surface glycoprotein MITat 1.6 - Trypanosoma brucei (fragment)

C;Species: Trypanosoma brucei

C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-May-1999

C;Accession: D61512

R;Holder: A.A.; Cross, G.A.M.

Mol. Biochem. Parasitol. 2, 135-150, 1981

A;Title: Glycopeptides from variant surface glycoproteins of Trypanosoma brucei. C-term

A;Reference number: A61512; MUID:81172836; PMID:6163983

A;Accession: D61512

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-8 <HOL>

C;Keywords: glycoprotein

Query Match 39.7%; Score 25; DB 2; Length 8;

Best Local Similarity 50.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 WSGHCC 8

Db 1 WSGETC 6

RESULT 2

A60522

sperm-activating peptide SAP-IV - sea urchin (Diadema setosum)

C;Species: Diadema setosum

C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000

C;Accession: A60522

R;Yoshino, K.I.; Kurita, M.; Yamaguchi, M.; Nomura, K.; Takao, T.; Shimonishi, Y.; Suzu

Comp. Biochem. Physiol. B 95, 423-429, 1990

A;Title: A species-specific sperm-activating peptide from the egg jelly of the sea urch

A;Reference number: A60522; MUID:90227916; PMID:2158412

A;Accession: A60522

A;Molecule type: protein

A;Residues: 1-9 <YOS>

C;Superfamily: unassigned animal peptides

F;2-9/Disulfide bonds: #status experimental

Query Match 36.5%; Score 23; DB 2; Length 9;

Best Local Similarity 50.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 WSGHCC 8

Db 4 WSGAVC 9

RESULT 3

C61512

variant surface glycoprotein MITat 1.4 - Trypanosoma brucei (fragment)

C:Species: Trypanosoma brucei
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-May-1999
C:Accession: C61512
R:Holder, A.A.; Cross, G.A.M.
Mol. Biochem. Parasitol. 2, 135-150, 1981
A:Title: Glycopeptides from variant surface glycoproteins of Trypanosoma brucei. C-terminal
A:Reference number: A61512; MUID:81172836; PMID:6163383
A:Accession: C61512
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <HOL>
C:Keywords: glycoprotein

Query Match 33.3%; Score 21; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 WSGHCC 8
| | |
| | |
DB 1 WENNAC 6

RESULT 4

PT0652
T-cell receptor beta chain V-D-J region (121-1E) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0652
R:Teeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0652
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <PEE>
A:Experimental source: day 4 postnatal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 28.6%; Score 18; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 SGHC 7
| | |
| | |
DB 2 SGDC 5

RESULT 5

I79564
Hypothetical TCL3 protein (mistranslated) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 28-Jan-2000 #sequence_revision 28-Jan-2000 #text_change 28-Jan-2000
C:Accession: I79564
R:Zutter, M.; Hockett, R.D.; Roberts, C.W.; McGuire, E.A.; Bloomstone, J.; Morton, C.C.;
Proc. Natl. Acad. Sci. U.S.A. 87, 3161-3165, 1990
A:Title: The t(10;14)(q24;q11) of T-cell acute lymphoblastic leukemia juxtaposes the del
A:Reference number: I59162; MUID:90222189; PMID:2326274
A:Accession: I79564
A>Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6 <ZUT>
A:Cross-references: GB:M33602; NID:g339907; PIDX:AAA66449.1; PID:g807656
C:Comment: This is the hypothetical translation of a sequence translated in an incorrect

Query Match 28.6%; Score 18; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CC 8
| | |
| | |
DB 3 CC 4

RESULT 6

PC1002
leucine-tRNA ligase (EC 6.1.1.4) - Escherichia coli (fragments)
N:Alternate names: leucyl-tRNA synthetase
C:Species: Escherichia coli
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 03-Jun-2002
C:Accession: PC1002
R:Miao, P.; Shi, J.P.; Wang, Y.L.
Science in China (series B) 34, 691-698, 1991
A:Title: Chemical modification of sulphydryl groups of E. coli leucyl-tRNA synthetase a
A:Reference number: PC1002
A:Accession: PC1002
A:Molecule type: protein
A:Residues: 1-8 <MIA>
C:Comment: This enzyme catalyzes the aminoacylation of tRNA(Leu) with Leucine.
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
P:5-8/Region: catalytic #status predicted

Query Match 28.6%; Score 18; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CC 8
| | |
| | |
DB 4 CC 5

RESULT 7

QDR8
delta sleep-inducing peptide - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 20-Jun-2000
C:Accession: A01422
R:Monnier, M.; Didier, L.; Gachter, R.; Maier, P.F.; Tobler, H.J.; Schoenenberger, G.A.
Experientia 33, 548-552, 1977
A:Title: The delta sleep inducing peptide (DSIP). Comparative properties of the origins
A:Reference number: A01422; MUID:77185324; PMID:862769
A:Accession: A01422
A:Molecule type: protein
A:Residues: 1-9 <MON>
C:Comment: This peptide was obtained from dialysates of occipital venous sinus blood fr
of recipient rabbits, it induces spindle and delta EEG activity and reduced motor acti
C:Superfamily: unassigned animal peptides

Query Match 28.6%; Score 18; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSG 5
| | |
| | |
DB 1 WAG 3

RESULT 8

A43848
cell surface adhesin for heparan sulfate, 66K - Staphylococcus aureus (fragment)
C:Species: Staphylococcus aureus
C>Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 24-Feb-1995
C:Accession: A43848
R:Liang, O.D.; Ascencio, F.; Fransson, L.A.; Wadstrom, T.
Infect. Immun. 60, 899-906, 1992
A:Title: Binding of heparan sulfate to Staphylococcus aureus.
A:Reference number: A43848; MUID:92176005; PMID:1541563
A:Accession: A43848
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <LIA>
A>Note: sequence extracted from NCBI backbone (NCBIP:85442)

Query Match 28.6%; Score 18; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;

Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSG 5
Db 2 WGG 4

RESULT 9

PT0562

T-cell receptor beta chain V-D-J region (126-1AK) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0562

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

A:Accession: PT0562

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-9 <FEE>

A:Experimental source: day 18 fetal thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 28.6%; Score 18; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSG 5
Db 6 NWGG 9

RESULT 10

PH0942

T-cell receptor beta chain V-D-J region (clone 13) - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997

C:Accession: PH0942

R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenberg, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991

A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic

A:Reference number: PH0891; MUID:92078857; PMID:1836012

A:Accession: PH0942

A:Molecule type: mRNA

A:Residues: 1-9 <GOL>

A:Experimental source: complete Freund's adjuvant-immunized lymph node

A:Note: the authors translated the codon TGC for residue 2 as Ala

C:Keywords: T-cell receptor

Query Match 28.6%; Score 18; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CC 8
Db 1 CC 2

RESULT 11

B53284

T-cell receptor beta 2 chain D region, Dbeta2 - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: B53284

R:Harindranath, N.; Alexander, C.B.; Mage, R.G.

Mol. Immunol. 28, 881-888, 1991

A:Title: Evolutionarily conserved organization and sequences of germline diversity and

A:Reference number: A53284; MUID:91342695; PMID:1678859

A:Accession: B53284

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-4 <HAR>

A:Cross-references: GB:S60737; NID:G233916; PIDN:AAB19518.1; PID:G233918

A:Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBI:P:60738)

C:Keywords: T-cell receptor

Query Match 27.0%; Score 17; DB 2; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WSG 5
Db 2 WGG 4

RESULT 12

S08606

hypothetical protein 2 estrogen receptor 5'-region - chicken

C:Species: Gallus gallus (chicken)

C>Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 18-Aug-2000

C:Accession: S08606

R:Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.; Bornert, J.M.; Chambon, P.

EMBO J. 5, 891-897, 1986

A:Title: The chicken oestrogen receptor sequence: homology with v-erbA and the human oe

A:Reference number: S07192; MUID:86247578; PMID:3755102

A:Accession: S08606

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-7 <FRU>

A:Cross-references: EMBL:X03805; NID:G63378; PIDN:CAA27432.1; PID:G584490

C:Superfamily: unassigned leader peptides

Query Match 27.0%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HC 7
Db 5 HC 6

RESULT 13

PT0642

T-cell receptor beta chain V-D-J region (111-1H) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0642

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

A:Accession: PT0642

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-7 <FEE>

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 27.0%; Score 17; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WSG 5
Db 5 WGG 7

RESULT 14

PT0586

T-cell receptor beta chain V-D-J region (141-1CN) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0586; PT0592

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0586
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <PEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c (clones 141-iCN and 141-ICD)
C;Keywords: T-cell receptor

Query Match 27.0%; Score 17; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. NO. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 WSG 5
|
Db 5 WGG 7

RESULT 15

PT0728
T-cell receptor beta chain V-D-J region (161-2H) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0728
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0728
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <PEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 27.0%; Score 17; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. NO. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 WSG 5
|
Db 5 WGG 7

Search completed: June 1, 2004, 10:28:46
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:25:45 ; Search time 11 Seconds
(without alignments)
42.603 Million cell updates/sec

Title: US-09-756-899A-1

Perfect score: 63

Sequence: 1 AHWSGHCL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 251

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	28.6	9	1 DSIP_RABIT	P01158 cryotolagus
2	17	27.0	9	1 TALI_PICJA	P17440 pichia jaci
3	17	27.0	9	1 TALI_PICJA	P17441 pichia jaci
4	16	25.4	9	1 COW CONVE	P83047 conus ventr
5	15.5	24.6	9	1 LIT0 LITAU	P08945 litoria aur
6	15.5	24.6	9	1 LITR_PHYRO	P08946 phylomedusa
7	15	23.8	5	1 UF01_MOUSE	P38639 mus musculus
8	15	23.8	9	1 NEP HV128	P12481 human immun
9	14	22.2	3	1 GRWY_HUMAN	P01157 homo sapien
10	14	22.2	4	1 DCM1_PSECH	P19916 pseudomonas
11	14	22.2	9	1 MOSH_CUYJA	P19852 clypeaster
12	14	22.2	9	1 NSK1_SARBU	P41492 sarcophaga
13	14	22.2	9	1 SAP_STOVA	P24047 stomopneute
14	13	20.6	5	1 BIO1_CITFR	P12997 citrobacter
15	13	20.6	9	1 DNFI_LOCFI	P16339 locusta mig
16	13	20.6	9	1 LMIP_LOCFI	P31799 locusta mig
17	13	20.6	9	1 PKPI_PPRAM	P82691 periplaneta
18	12	19.0	5	1 BPP7_BOTIN	P30425 bothrops in
19	12	19.0	6	1 TRPI_PSEPU	P36414 pseudomonas
20	12	19.0	8	1 COW2_CONFU	P58785 conus purpu
21	12	19.0	8	1 HTF1_PPRAM	P04548 periplaneta
22	12	19.0	8	1 HTF2_PPRAM	P04549 periplaneta
23	12	19.0	8	1 HTP_TENNO	P25419 tenebrio mo
24	12	19.0	9	1 XYL4_STRSQ	P19149 streptomyce
25	11	17.5	4	1 OCP3_OCTMI	P58649 octopus min
26	11	17.5	6	1 E101_LITRU	P82096 litoria rub
27	11	17.5	6	1 LOK1_LOCFI	P41491 locusta mig
28	11	17.5	7	1 BRHP_CONIM	P58803 conus imper
29	11	17.5	7	1 TPFY_PACDA	P83455 pachymedusa
30	11	17.5	7	1 TY51_LITRU	P82065 litoria rub
31	11	17.5	7	1 WWA1_ACHFU	P35919 achatina fu
32	11	17.5	7	1 WWA2_ACHFU	P35920 achatina fu
33	11	17.5	7	1 WWA3_ACHFU	P35921 achatina fu

```

34 11 17.5 8 1 AC1_THUAL      P18691 thunnus alb
35 11 17.5 8 1 AKHG_GRYBI     P14086 gryllus bim
36 11 17.5 8 1 AKH_LIBAU     P25418 libellula a
37 11 17.5 8 1 AKH_WELML     P25423 melolontha
38 11 17.5 8 1 AKH_TABAT     P14595 tabanus atr
39 11 17.5 8 1 ALI1_CYDPO     P82152 cydia pomon
40 11 17.5 8 1 CCKN_MACRU     P30369 macropus eu
41 11 17.5 8 1 LCK1_LEUMA     P21140 leucophaea
42 11 17.5 8 1 LCK2_LEUMA     P21141 leucophaea
43 11 17.5 8 1 LCK3_LEUMA     P21142 leucophaea
44 11 17.5 8 1 LCK4_LEUMA     P21143 leucophaea
45 11 17.5 8 1 LCK5_LEUMA     P19987 leucophaea

```

ALIGNMENTS

```

RESULT 1
DSIP_RABIT
ID_DSIP_RABIT STANDARD; PRT; 9 AA.
AC P01158;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Delta sleep-inducing peptide (DSIP).
OS Cryotolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=77185324; PubMed=862769;
RA Monnier M., Dudler L., Gachter R., Maier P.F., Tobler H.J.,
RA Schoenenberger G.A.;
RT "The delta sleep inducing peptide (DSIP). Comparative properties of
RT the original and synthetic nonapeptide.";
RL Expientia 33:548-552(1977).
RN [2]
RP SEQUENCE, AND SYNTHESIS.
RX MEDLINE=79054421; PubMed=568769;
RA Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Monnier M.;
RT "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino-acid
RT analysis, sequence, synthesis and activity of the nonapeptide.";
RN [3]
RP FLUGERS ARCH. 376:119-129(1978).
RN [3]
RP REVIEW.
RX MEDLINE=87175129; PubMed=3550726;
RA Graf M.V., Kestin A.J.;
RT "Delta-sleep-inducing peptide (DSIP): an update.";
RL Peptides 7:1165-1187(1986).
CC -!- FUNCTION: When infused into the mesodiencephalic ventricle of
CC recipient rabbits induces spindle and delta EEG activity and
CC reduced motor activities.
CC -!- MISCELLANEOUS: This peptide was obtained from dialysates of
CC occipital venous sinus blood from rabbits kept asleep by electric
CC stimulation of the thalamus.
CC -!- DATABASE: NAME=Protein Spotlight;
CC NOTE=Issue 8 of March 2001;
CC WWW=http://www.expasy.org/spotlight/articles/splt008.html".
DR PIR: A01422; QDRB.
SQ SEQUENCE 9 AA; 849 MW; DDD365BDDAA8787D CRC64;

```

Query Match 28.6%; Score 18; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSG 5
DB 1 WAG 3

RESULT 2
TALI_PICJA

```

ID TAL1 PICUA STANDARD; PRT; 9 AA.
AC P17410;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transaldolase I (EC 2.2.1.2) (Fragment).
OS Pichia jadinii (Yeast) (Candida utilis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4903;
RN [1]
RP SEQUENCE
RX MEDLINE=71110646; PubMed=556924;
RA Sun S.C., Goris L., Tsolas O.;
RT "Purification of crystallization of transaldolase isozyme I and
RT evidence for different genetic origin of isozymes I and III in
RT Candida utilis.";
RL Arch. Biochem. Biophys. 178:69-78(1977).
CC -!- FUNCTION: Transaldolase is important for the balance of
CC metabolites in the pentose-phosphate pathway.
CC -!- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde
CC 3-phosphate = D-erythrose 4-phosphate + D-fructose 6-phosphate.
CC -!- PATHWAY: Pentose phosphate pathway; nonoxidative part.
CC -!- SIMILARITY: Belongs to the transaldolase family. Subfamily 1.
DR PIR; A12872; A12872.
DR InterPro; IPR001585; Transaldolase.
DR PROSITE; PS00958; TRANSALDOLASE_2; PARTIAL.
DR PROSITE; PS01054; TRANSALDOLASE_1; PARTIAL.
KW Transferase; Pentose shunt.
FT NON_TER 1
FT NON_TER 9
FT NON_TER 9
SQ SEQUENCE 9 AA; 1008 MW; 274F31AF0EB1E058 CRC64;

Query Match 27.0%; Score 17; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HC 7
DB 4 HC 5

RESULT 3
ID TAL3 PICUA STANDARD; PRT; 9 AA.
AC P17411;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transaldolase III (EC 2.2.1.2) (Fragment).
OS Pichia jadinii (Yeast) (Candida utilis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4903;
RN [1]
RP SEQUENCE
RX MEDLINE=75145197; PubMed=1092268;
RA Tsolas O., Sun S.C.;
RT "Isolation of a peptide containing a histidyl-cysteiny sequence
RT from the active center of transaldolase.";
RL Arch. Biochem. Biophys. 167:525-533(1975).
CC -!- FUNCTION: Transaldolase is important for the balance of
CC metabolites in the pentose-phosphate pathway.
CC -!- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde
CC 3-phosphate = D-erythrose 4-phosphate + D-fructose 6-phosphate.
CC -!- PATHWAY: Pentose phosphate pathway; nonoxidative part.
CC -!- SIMILARITY: Belongs to the transaldolase family. Subfamily 1.
DR PIR; A11497; A11497.
DR InterPro; IPR001585; Transaldolase.
DR PROSITE; PS00958; TRANSALDOLASE_2; PARTIAL.
DR PROSITE; PS01054; TRANSALDOLASE_1; PARTIAL.
KW Transferase; Pentose shunt.
FT NON_TER 1
FT NON_TER 1
```

```

FT NON_TER 9
SQ SEQUENCE 9 AA; 1033 MW; 325A31A44EB1E058 CRC64;

Query Match 27.0%; Score 17; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HC 7
DB 4 HC 5

RESULT 4
ID COW CONVE STANDARD; PRT; 9 AA.
AC P83047;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Contryphan-Vn.
OS Conus ventricosus (Mediterranean cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeconcha; Hypogastropoda;
OC Negastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=117992;
RN [1]
RP SEQUENCE, SYNTHESIS, DISULFIDE BONDS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=21547785; PubMed=11689995;
RA Massilia G.R., Schinina M.E., Ascenzi P., Polticelli F.;
RT "Contryphan-Vn: a novel peptide from the venom of the Mediterranean
RT snail Conus ventricosus.";
RL Biochem. Biophys. Res. Commun. 288:908-913(2001).
RN [2]
RP STRUCTURE BY NMR, SYNTHESIS, DISULFIDE BONDS, AND FUNCTION.
RX MEDLINE=22533339; PubMed=12646193;
RA Massilia G.R., Elisei V., Grolleau F., Lapiet B., Barbier J.,
RA Bournaud R., Molgo J., Cicero D.O., Paci M., Schinina M.E.,
RA Ascenzi P., Polticelli F.;
RT "Contryphan-Vn: a modulator of Ca2+-dependent K+ channels.";
RL Biochem. Biophys. Res. Commun. 303:238-246(2003).
CC -!- FUNCTION: Affects both voltage-gated and calcium-dependent
CC potassium channel activities, with composite and diversified
CC effects in invertebrate and vertebrate systems.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- PTM: The cis isomer is the most abundant and is thus thought to be
CC the functionally relevant conformer.
CC -!- MASS SPECTROMETRY: MW=1088.6; METHOD=MALDI.
CC -!- SIMILARITY: Belongs to the contryphan family.
DR PDB; 1NXX; 04-MAR-03.
KW Toxin; Ionic channel inhibitor; Neurotoxin;
KW Potassium channel inhibitor; D-amino acid; Amidation; 3D-structure.
FT DISULFID 3
FT MOD_RES 5 5 D-TRYPTOPHAN.
FT MOD_RES 9 9 AMIDATION.
FT MOD_RES 9 9
SQ SEQUENCE 9 AA; 1091 MW; 8D38676323676EBA CRC64;

Query Match 25.4%; Score 16; DB 1; Length 9;
Best Local Similarity 40.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 WSGHC 7
DB 5 WKFWC 9

RESULT 5
ID LITO LITAU STANDARD; PRT; 9 AA.
AC P089Z5;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
```

DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Litorin.
OS Litoria aurea (Green and golden bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;
OC Peleodyadinae; Litoria.
OX NCBi_TaxID=8371;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=75187011; PubMed=1140241;
RA Anastasi A., Erspamer V., Endean R.;
RT "Amino acid composition and sequence of litorin, a bombesin-like
nonapeptide from the skin of the Australian leptodactylid frog
Litoria aurea.";
RL Experientia 31:510-511(1975).
RN [2]
RP SEQUENCE, AND METHYLATION OF GLN-2.
RC TISSUE=Skin secretion;
RX MEDLINE=78003546; PubMed=908397;
RA Anastasi A., Montecucchi P.C., Angelucci F., Erspamer V., Endean R.;
RT "Glu(OMe)3-litorin, the second bombesin-like peptide occurring in
methanol extracts of the skin of the Australian frog Litoria aurea.";
RL Experientia 33:1289-1289(1977).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/ratanensin
family.
DR PIR: S07204; S07204.
DR PIR: S07205; S07205.
DR InterPro: IPR000874; Bombesin.
DR Pfam: PF02044; Bombesin; 1.
DR PROSITE: PS00257; BOMBESIN; 1.
KW Amphibian defense peptide; Bombesin family; Amidation;
KW Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 2 2 DEAMIDATION AND METHYLATION (PARTIAL).
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1103 MW; D7CCC1B862CDC366 CRC64;
Query Match 24.6%; Score 15.5; DB 1; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 3 WS-GH 6
DB 3 WAVGH 7
RESULT 6
LITR_PHYTO
ID_LITR_PHYTO STANDARD; PRT; 9 AA.
AC P08946;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DS Rhodol-litorin.
OS Phyllomedusa rohdei (Rohde's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;
OC Phyllomedusinae; Phyllomedusa.
OX NCBi_TaxID=8394;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=85127560; PubMed=3838283;
RA Barra D., Erspamer G.F., Simmaco M., Bossa F., Melchiorri P.,
RA Erspamer V.;
RT "Rhodol-litorin: a new peptide from the skin of Phyllomedusa rohdei.";
RL FEBS Lett. 182:53-55(1985).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Belongs to the bombesin/neuromedin B/ratanensin

CC family.
DR PIR: S07241; S07241.
DR InterPro: IPR000874; Bombesin.
DR Pfam: PF02044; Bombesin; 1.
DR PROSITE: PS00257; BOMBESIN; 1.
KW Amphibian defense peptide; Bombesin family; Amidation;
KW Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1090 MW; 4ECCC1E861ADC377 CRC64;
Query Match 24.6%; Score 15.5; DB 1; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 3 W-SGH 6
DB 3 WAVGH 7
RESULT 7
UF01_MOUSE
ID UF01_MOUSE STANDARD; PRT; 5 AA.
AC P38639;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page of fibroblasts (F19) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBi_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Fibroblast;
RX MEDLINE=95009907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins
RT using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
protein is: 6.6, its MW is: 19 kDa.
FT NON_TER 5
SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;
Query Match 23.8%; Score 15; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 WSG 5
DB 1 WIG 3
RESULT 8
NEF_HV128
ID_NEF_HV128 STANDARD; PRT; 9 AA.
AC P12481;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (P-protein) (27 kDa protein) (3'ORF) (Fragment).
GN NEF.
OS Human immunodeficiency virus type 1 (2-84 isolate) (HIV-1).
OC Viruses; Retroviruses; Retroviridae; Lentivirus.
OX NCBi_TaxID=11681;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88281279; PubMed=3395517;
RA Younso J., Josephs S.P., Reitz M.S. Jr., Zagury D., Wong-Staal F.,
RA Gallo R.C.;
RT "Nucleotide sequence analysis of the env gene of a new Zairian
RT isolate of HIV-1.";

```

AIDS Res. Hum. Retroviruses 4:165-173(1988).
-!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -!- MISCELLANEOUS: THE Z-84 ISOLATE WAS TAKEN FROM A 54 YEAR-OLD
CC ZAISEAN MALE.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03653; AAA44687.1; -.
CC DR HIV; J03653; NEFSJYL.
CC DR AIDS; Myristate; GTP-binding; Lipoprotein.
CC LIPID 2 2 N-myristoyl glycine (in host) (By
CC similarity).
CC FT NON_TER 9 9
CC FT SEQUENCE 9 AA; 967 MW; 319CB325A3733878 CRC64;
CC
CC Query Match 23.8%; Score 15; DB 1; Length 9;
CC Best Local Similarity 100.0%; Pred. No. 1.4e+05;
CC Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Qy 3 WS 4
CC |||
CC DB 5 WS 6

```

```

RESULT 9
GRNW_HUMAN
ID GRNW_HUMAN STANDARD; PRT; 3 AA.
FOI157;
AC AC
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Growth-modulating peptide.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RN
RP SEQUENCE.
MEDLINE=77162359; PubMed=859356;
RA Schlesinger D.B., Pickart L., Thaler M.M.;
RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine.";
RL Experimentia 33:324-325(1977).
-1- MISCELLANEOUS: This serum tripeptide has been found to stimulate
CC growth of some cell types and to inhibit other types in vitro.
CC GO: 0001558; P: regulation of cell growth; NAC.
CC GO: 0001558; P: regulation of cell growth; NAC.
DR GO SEQUENCE 3 AA; 340 MW; 633188100000000 CRC64;
SQ

```

```

Query Match      22.2%; Score 14; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 GH 6
        ||
Db      1 GH 2

RESULT 10
DCML PSECH
ID _DCML_PSECH STANDARD; PRT; 4 AA.
AC P19916;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
DE dehydrogenase subunit L) (CO-DH L) (Fragment).
GN CUTL.

```

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GNC 7
DB 6 GNC 8

RESULT 14
BIOS_CITER
ID BIOS_CITER STANDARD; PRT; 5 AA.
AC P12997;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Biotin synthase (EC 2.8.1.6) (Biotin synthetase) (Fragment).
GN BIOS.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN [1]
RP SEQUENCE FROM N.A. PubMed=2971595;
RX MEDLINE=89006280; PubMed=2971595;
RA Shiuan D., Campbell A.;
RT "Transcriptional regulation and gene arrangement of Escherichia coli,
Citrobacter freundii and Salmonella typhimurium biotin operons."
RL Gene 67:203-211(1988).
CC -!- CATALYTIC ACTIVITY: Dethiobiotin + sulfop. = biotin.
CC -!- PATHWAY: Biotin biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the biotin and lipoic acid synthetases family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; M21922; -; NOT_ANNOTATED_CDS.
CC DR EIR; 140698; 140698.
KW Biotin biosynthesis; Iron-sulfur; Transferase.
FT NON_TER 5
SQ SEQUENCE 5 AA; 532 MW; 75A5E1BDD6F00000 CRC64;

Query Match 22.2%; Score 14; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GH 6
DB 5 GH 6

RESULT 13
SAP_STOVA
ID SAP_STOVA STANDARD; PRT; 9 AA.
AC P24047;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sperm-activating peptide (SAP).
OS Stomopneustes variolaris (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidae; Diademataceae; Phymosomatoida; Stomechinidae;
OC Stomopneustes.
OX NCBI_TaxID=7663;
RN [1]
RP SEQUENCE, AND DISULFIDE BOND.
RC TISSUE=Egg jelly;
RX MEDLINE=92097763; PubMed=1756858;
RA Yoshino K.-I., Takao T., Shimomishi Y., Suzuki N.;
RT "Determination of the amino acid sequence of an intramolecular
RT disulfide linkage-containing sperm-activating peptide by tandem mass
RT spectrometry."
RL FEBS Lett. 294:179-182(1991).
CC -!- FUNCTION: Cause stimulation of sperm respiration and motility
CC through intracellular alkalization, transient elevations of
CC cAMP, cGMP and calcium levels in sperm cells, and transient
CC activation and subsequent inactivation of the membrane form of
CC Guanylate cyclase.
FT DISULFID 3
SQ SEQUENCE 9 AA; 1010 MW; C469B3387B076E9 CRC64;

Query Match 22.2%; Score 14; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;

QY 5 GH 6
DB 5 GH 6

RESULT 15
DNFI_LOCHI
ID DNFI_LOCHI STANDARD; PRT; 9 AA.
AC P16339;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Locupresin (Diuretic neuropeptide F1/F2).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Subesophageal ganglion, and Thoracic ganglion;
RX MEDLINE=88077077; PubMed=3689410;
RA Proux J.P., Miller C.A., Li J.P., Carney R.L., Girardie A.,

RESULT 12
NSKI_SARBU
ID NSKI_SARBU STANDARD; PRT; 9 AA.
AC P41452;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neosulfakinin-I (NEB-SK-I).
OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Sarcophagidae; Sarcophaga.
OX NCBI_TaxID=7385;
RN [1]
RP SEQUENCE.
RC TISSUE=Head;
RX MEDLINE=93083101; PubMed=1360367;
RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.;
RT "Isolation and primary structure of two sulfakinin-like peptides from
RT the fleshfly, Neobellieria bullata."
RL Comp. Biochem. Physiol. 103C:135-142(1992).
CC -!- FUNCTION: Myotropic peptide.
CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; 1.
KW Neuropeptide; Amidation; Sulfation.
FT MOD_RES 4 4 SULFATION (POTENTIAL).
FT MOD_RES 9 9 AMIDATION (POTENTIAL).
SQ SEQUENCE 9 AA; 1187 MW; 8B0A0691E86B5AAA CRC64;

Query Match 22.2%; Score 14; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GH 6
DB 5 GH 6

RESULT 13
SAP_STOVA
ID SAP_STOVA STANDARD; PRT; 9 AA.
AC P24047;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sperm-activating peptide (SAP).
OS Stomopneustes variolaris (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidae; Diademataceae; Phymosomatoida; Stomechinidae;
OC Stomopneustes.
OX NCBI_TaxID=7663;
RN [1]
RP SEQUENCE, AND DISULFIDE BOND.
RC TISSUE=Egg jelly;
RX MEDLINE=92097763; PubMed=1756858;
RA Yoshino K.-I., Takao T., Shimomishi Y., Suzuki N.;
RT "Determination of the amino acid sequence of an intramolecular
RT disulfide linkage-containing sperm-activating peptide by tandem mass
RT spectrometry."
RL FEBS Lett. 294:179-182(1991).
CC -!- FUNCTION: Cause stimulation of sperm respiration and motility
CC through intracellular alkalization, transient elevations of
CC cAMP, cGMP and calcium levels in sperm cells, and transient
CC activation and subsequent inactivation of the membrane form of
CC Guanylate cyclase.
FT DISULFID 3
SQ SEQUENCE 9 AA; 1010 MW; C469B3387B076E9 CRC64;

Query Match 22.2%; Score 14; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.4e+05;

QY 5 GH 6
DB 5 GH 6

RESULT 15
DNFI_LOCHI
ID DNFI_LOCHI STANDARD; PRT; 9 AA.
AC P16339;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Locupresin (Diuretic neuropeptide F1/F2).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Subesophageal ganglion, and Thoracic ganglion;
RX MEDLINE=88077077; PubMed=3689410;
RA Proux J.P., Miller C.A., Li J.P., Carney R.L., Girardie A.,

RA Delaage M., Schooley D.A.;
RT "Identification of an arginine vasopressin-like diuretic hormone from
RT Locusta migratoria",
RL Biochem. Biophys. Res. Commun. 149:180-186(1987).
CC -!- FUNCTION: DIURETIC HORMONE.
CC -!- SUBUNIT: Monomer (F1) and homodimer (F2); disulfide-linked.
CC -!- SIMILARITY: Belongs to the vasopressin/oxytocin family.
DR PIR: A29477; A29477.
DR InterPro: IPR000981; Neurhyp_horm.
DR Pfam: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Neuropeptide; Amidation.
FT DISULFID 1 6 IN FI.
FT DISULFID 1 1 INTERCHAIN (WITH C-6) (IN F2).
FT DISULFID 6 6 INTERCHAIN (WITH C-1) (IN F2).
FT MOD RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 976 MW; 56EB176EB451A057 CRC64;

Query Match 20.6%; Score 13; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 CL 9
Db 1 CL 2

Search completed: June 1, 2004, 10:27:20
Job time : 13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:25:46 ; Search time 39 Seconds
(without alignments)
72.812 Million cell updates/sec

Title: US-09-756-899A-1
Perfect score: 63
Sequence: 1 AHWGHCCL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 840

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_todent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	34.9	9	10 Q9FSZ2	Q9fsz2 cicer ariet
2	19.5	31.0	8	6 O02831	O02831 oryctolagus
3	19	30.2	8	4 Q8IV87	Q8iv87 homo sapien
4	19	30.2	9	8 Q8W8X4	Q8w8x4 diadema mex
5	18	28.6	8	13 Q98TU5	Q98tu5 xenopus lae
6	18	28.6	9	2 Q9S5M1	Q9s5m1 staphylococ
7	18	28.6	9	11 Q08979	Q08979 mus musculu
8	18	28.6	9	12 Q9E1U7	Q9e1u7 hepatitis b
9	17	27.0	8	4 Q15900	Q15900 homo sapien
10	17	27.0	9	4 Q99887	Q99887 homo sapien
11	17	27.0	9	15 Q85723	Q85723 simian sarc
12	16	25.4	9	8 Q94NB1	Q94nb1 microcebus
13	16	25.4	9	8 Q94NB2	Q94nb2 microcebus
14	16	25.4	9	8 Q94NA9	Q94na9 daubentonla
15	16	25.4	9	8 Q94XE6	Q94xe6 tectocoris
16	16	25.4	9	8 Q94NB0	Q94nb0 microcebus

17	16	25.4	9	8 Q85DB8	Q85db8 lepilemur e
18	16	25.4	9	8 Q85DB0	Q85db0 lepilemur s
19	15	23.8	8	4 Q15888	Q15888 homo sapien
20	15	23.8	8	8 Q94VC1	Q94vc1 varanus rud
21	15	23.8	8	8 Q9TD02	Q9td02 terranatos
22	15	23.8	8	8 Q94PX5	Q94px5 felis silve
23	15	23.8	8	8 Q94VB2	Q94vb2 varanus sal
24	15	23.8	8	8 Q94PX7	Q94px7 felis silve
25	15	23.8	8	8 Q94PX6	Q94px6 felis libyc
26	15	23.8	8	8 Q94VA7	Q94va7 varanus sal
27	15	23.8	8	8 Q94VB5	Q94vb5 varanus sal
28	15	23.8	8	11 P82598	P82598 rattus norv
29	15	23.8	8	12 Q64971	Q64971 alfalfa mos
30	15	23.8	9	2 Q47410	Q47410 escherichia
31	15	23.8	9	4 Q9UC36	Q9uc36 homo sapien
32	15	23.8	9	8 Q94VH4	Q94vh4 varanus gla
33	15	23.8	9	8 Q94V18	Q94v18 varanus ere
34	15	23.8	9	8 Q94VC6	Q94vc6 varanus pil
35	15	23.8	9	13 Q9PRJ4	Q9prj4 lepisosteus
36	14	22.2	7	10 P82445	P82445 nicotiana t
37	14	22.2	8	6 Q9TRY3	Q9try3 sus sp. ins
38	14	22.2	9	4 Q15999	Q15999 homo sapien
39	14	22.2	9	12 Q89491	Q89491 murine minu
40	13	20.6	7	12 Q66113	Q66113 cherry leaf
41	13	20.6	7	13 Q42564	Q42564 fugu rubrip
42	13	20.6	8	2 Q85406	Q85406 coxiella bu
43	13	20.6	9	2 Q47063	Q47063 escherichia
44	13	20.6	9	4 Q9UMF3	Q9umf3 homo sapien
45	13	20.6	9	11 Q61723	Q61723 mus musculu

ALIGNMENTS

RESULT 1

Q9FSZ2 PRELIMINARY; PRT; 9 AA.
AC Q9FSZ2 (TREMREL. 16, Created)
DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Cicer arietinum (Chickpea) (Garbanzo).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
CX NCBI_TaxID=3827;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-cv. Castellana; TISSUE=Etisolated epicotyl;
RC Dopico B., Jimenez T., Labrador E.;
RA "cDNA clones expressed in etiolated Cicer arietinum epicotyls.";
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ299069; CAC10216.1; -
DR EMBL; Hypothetical protein.
KW NON TER
FT NON TER
SQ SEQUENCE 9 AA; 990 MW; 9441BDDAA7272EBB CRC64;

Query Match 34.9%; Score 22; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 16-06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 CCL 9
Db 2 CCL 4

RESULT 2

O02831 PRELIMINARY; PRT; 8 AA.
ID O02831
AC O02831
DT 01-JUL-1997 (TREMREL. 04, Created)
DT 01-JUL-1997 (TREMREL. 04, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Pro alpha 1 type III collagen protein (fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96377333; PubMed=8783186;
 RA Metsarantha M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
 RA Vuorio E.;
 RT "Evidence for insufficient chondrocytic differentiation during repair
 RT of full-thickness defects of articular cartilage.";
 RL Matrix Biol. 15:39-47(1996).
 DR EMBL; S83371; AAD14433.1; -.
 KW Collagen.
 FT NON_TER
 SQ SEQUENCE 8 AA; 1028 MW; B859C7272EA77371 CRC64;

 Query Match 31.0%; Score 19.5; DB 6; Length 8;
 Best Local Similarity 50.0%; Pred. No. 1e+06;
 Matches 4; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

 QY 2 HWSGHCLL 9
 DB 1 HWP---CL 5

 RESULT 3
 Q8IV87 PRELIMINARY; PRT; 8 AA.
 AC Q8IV87;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE DJ107712.2 (Serine palmitoyltransferase, long chain base subunit
 D3 2-like (Aminotransferase 2), variant 1) (Fragment).
 GN SPPLC2L.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith M.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL050320; CAD54807.1; -.
 DR GO; GO:0016740; F:transferase activity; IEA.
 KW Transferase.
 FT NON_TER
 SQ SEQUENCE 8 AA; 908 MW; 8E533682CEBEB042 CRC64;

 Query Match 30.2%; Score 19; DB 4; Length 8;
 Best Local Similarity 66.7%; Pred. No. 1e+06;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 7 CCL 9
 DB 3 CCV 5

 RESULT 4
 Q8W8X4 PRELIMINARY; PRT; 9 AA.
 AC Q8W8X4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytochrome oxidase subunit II (fragment).
 GN COII.
 OS Diadema mexicanum.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

OC Echinoidea; Euechinoidea; Diademataceae; Diadematoidea; Diadematiidae;
 OC Diadema.
 OX NCBI_TaxID=105359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CC70, and CC117;
 RX MEDLINE=21323357; PubMed=11430656;
 RA Lessios H.A., Kessing B.D., Pearse J.S.;
 RT "Population structure and speciation in tropical seas: global
 RT phylogeography of the sea urchin *Diadema*.";
 RL Evolution 55:955-975(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CC70, and CC117;
 RX MEDLINE=21561594; PubMed=11703875;
 RA Lessios H.A., Garrido M.J., Kessing B.D.;
 RT "Demographic history of *Diadema antillarum*, a keystone herbivore on
 RT Caribbean reefs.";
 RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
 DR EMBL; AY012920; AAL3843.1; -.
 DR EMBL; AY012921; AAL3844.1; -.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON_TER
 SQ SEQUENCE 9 AA; 1174 MW; 2B73173B46DDC2D3 CRC64;

 Query Match 30.2%; Score 19; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1e+06;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 HW 3
 DB 1 HW 2

 RESULT 5
 Q98TU5 PRELIMINARY; PRT; 8 AA.
 AC Q98TU5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ElrD transcript 2 (fragment).
 GN ELRD.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21226152; PubMed=11327714;
 RA Nassar F., Wegnez M.;
 RT "Characterization of two promoters of the *Xenopus laevis* elrD gene.";
 RL Biochem. Biophys. Res. Commun. 283:392-398(2001).
 DR EMBL; AF329448; AAK01428.1; -.
 FT NON_TER
 SQ SEQUENCE 8 AA; 1008 MW; FF93372874537B16 CRC64;

 Query Match 28.6%; Score 18; DB 13; Length 8;
 Best Local Similarity 66.7%; Pred. No. 1e+06;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 3 WSG 5
 DB 3 WNG 5

 RESULT 6
 Q9R5M1 PRELIMINARY; PRT; 9 AA.
 ID Q9R5M1
 AC Q9R5M1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE 66 kDa cell surface adhesin for heparan sulfate (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacilliales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE
RX MEDLINE=92176005; PubMed=1541563;
RA Ljung O.D., Ascencio F., Fransson L.A., Wadstrom T.;
RT "Binding of heparan sulfate to Staphylococcus aureus.";
RL Infect. Immun. 60:899-906(1992).
DR PIR; A43848; A43848.
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 990 MW; 2289DD7337861B3 CRC64;

Query Match 28.6%; Score 18; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSG 5
DB 2 WTG 4

RESULT 7
Q08979 ID Q08979 PRELIMINARY; PRT; 9 AA.
AC Q08979;
DT 01-JUL-1997 (T-EMBLrel. 04, Created)
DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE AML1 protein (Fragment).
GN AML1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Tumor;
RX MEDLINE=97332339; PubMed=9188573;
RA Antoft H.W., Sorensen A.B., Bareil C., Schmidt J., Luz A.,
RA Pedersen F.S.;
RT "Stability of AML1 (core) site enhancer mutations in T-lymphomas
RT induced by attenuated SL3-3 murine leukemia virus mutants.";
RL J. Virol. 71:5080-5087(1997).
DR EMBL; Y11802; CAA72496.1; -.
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 981 MW; 293E01E865A776DB CRC64;

Query Match 28.6%; Score 18; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SGH 6
DB 5 SGH 7

RESULT 8
Q9E1U7 ID Q9E1U7 PRELIMINARY; PRT; 9 AA.
AC Q9E1U7;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE X protein (Fragment).
GN X.
OS Hepatitis B virus.
OC Viruses; Retroviridae; Orthohepadnavirus.

OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21213459; PubMed=11315638;
RA Ha-Lee Y.M., Lee J., Pyun H., Kim Y., Kim Y.;
RT "Sequence variation of Hepatitis B virus promoter regions in
RT persistently infected patients.";
RL Arch. Virol. 146:279-292(2001).
DR EMBL; AF276526; AAG29993.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 994 MW; CD0FDEBEA2D40DD CRC64;

Query Match 28.6%; Score 18; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CC 8
DB 6 CC 7

RESULT 9
Q15900 ID Q15900 PRELIMINARY; PRT; 8 AA.
AC Q15900;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE (Clone XP7B11A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yardeni A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsey E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32079; AAA73890.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 931 MW; B5DDC40369AAEB1 CRC64;

Query Match 27.0%; Score 17; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HC 7
DB 1 HC 2

RESULT 10
Q99887 ID Q99887 PRELIMINARY; PRT; 9 AA.
AC Q99887;
DT 01-MAY-1997 (T-EMBLrel. 03, Created)
DT 01-MAY-1997 (T-EMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE 11 <beta>-HSD2 protein (Fragment).
GN 11 <BETA>-HSD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96133030; PubMed=8538347;
```

RA Stewart P.M., Krozowski Z.S., Gupta A., Milford D.V., Howie A.J.,
 RA Sheppard M.C., Whorwood C.B.;
 RT "Hypertension in the syndrome of apparent mineralocorticoid excess due
 to mutation of the 11 beta-hydroxysteroid dehydrogenase type 2 gene.";
 RL Lancet 347:88-91(1996).
 RL EMBL; S80133; RAD14324.1; -
 DR GO; GO:0003845; F11-beta-hydroxysteroid dehydrogenase activity; NAS.
 DR GO; GO:0008212; P:iminoalcoholcorticoid metabolism; NAS.
 FT NON TER 1
 SQ SEQUENCE 9 AA; 1020 MW; C3FC2BB1F5B059C9 CRC64;

Query Match 27.0%; Score 17; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1e+06; 0; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0;

Qy 6 HC 7
 ||
 Db 6 HC 7

RESULT 11
 Q85723 PRELIMINARY; PRT; 9 AA.
 ID Q85723;
 AC Q85723;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE V-sis and p28-sis genes (Fragment).
 OS Simian sarcoma virus.
 OC Viruses; Retrod viruses; Retroviridae; Gammaretrovirus.
 OC NCBI_TaxID=11817;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84106822; PubMed=6319011;
 RA Devare S.G., Shatzman A., Robbins K.C., Rosenberg M., Aaronson S.A.;
 RT "Expression of the PDGF-related transforming protein of simian sarcoma
 virus in E. coli.";
 RL Cell 36:43-49(1984).
 RL EMBL; K01473; AAA46816.1; -
 FT NON TER 9
 SQ SEQUENCE 9 AA; 1048 MW; 9C53A866C361A731 CRC64;

Query Match 27.0%; Score 17; DB 15; Length 9;
 Best Local Similarity 66.7%; Pred. No. 1e+06; 1; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 1;

Qy 3 MSG 5
 ||
 Db 5 MQG 7

RESULT 12
 Q94NB1 PRELIMINARY; PRT; 9 AA.
 ID Q94NB1;
 AC Q94NB1;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Cytochrome oxidase subunit III (Fragment).
 GN COIII.
 OS Microcebus ravelobensis.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
 OC Microcebus.
 OC NCBI_TaxID=122231;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=JP299, and JP301;
 RX MEDLINE=21184272; PubMed=11286490;
 RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
 RT "Molecular phylogeny of the lemur family Cheirogaleidae (Primates)
 based on mitochondrial DNA sequences.";

RL Mol. Phylogenet. Evol. 19:45-56(2001).
 DR EMBL; AF224630; AAK70571.1; -
 DR EMBL; AF224631; AAK70575.1; -
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON TER 1
 SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 25.4%; Score 16; DB 8; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1e+06; 1; Indels 0; Gaps 0;
 Matches 2; Conservative 1; Mismatches 1;

Qy 2 HMSG 5
 :||
 Db 5 YWVG 8

RESULT 13
 Q94NB2 PRELIMINARY; PRT; 9 AA.
 ID Q94NB2;
 AC Q94NB2;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Cytochrome oxidase subunit III (Fragment).
 GN COIII.
 OS Microcebus murinus (Lesser mouse lemur).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
 OC Microcebus.
 OC NCBI_TaxID=30608;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=JP285, JP288, JP289, JP292, JP308, and JP313;
 RX MEDLINE=21184272; PubMed=11286490;
 RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
 RT "Molecular phylogeny of the lemur family Cheirogaleidae (Primates)
 based on mitochondrial DNA sequences.";
 RL Mol. Phylogenet. Evol. 19:45-56(2001).
 DR EMBL; AF224624; AAK70547.1; -
 DR EMBL; AF224625; AAK70551.1; -
 DR EMBL; AF224626; AAK70555.1; -
 DR EMBL; AF224627; AAK70559.1; -
 DR EMBL; AF224628; AAK70563.1; -
 DR EMBL; AF224629; AAK70567.1; -
 DR GO; GO:0005739; C:mitochondrion; IEA.
 KW Mitochondrion.
 FT NON TER 1
 SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 25.4%; Score 16; DB 8; Length 9;
 Best Local Similarity 50.0%; Pred. No. 1e+06; 1; Indels 0; Gaps 0;
 Matches 2; Conservative 1; Mismatches 1;

Qy 2 HMSG 5
 :||
 Db 5 YWVG 8

RESULT 14
 Q94NA9 PRELIMINARY; PRT; 9 AA.
 ID Q94NA9;
 AC Q94NA9;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Cytochrome oxidase subunit III (Fragment).
 GN COIII.
 OS Daubentonia madagascariensis (Aye-aye).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Strepsirhini; Daubentonidae;

OC Daubentonina.
OX NCBI_TaxID=31869;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JP119, and JP120;
RX MEDLINE=21184272; PubMed=11286490;
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT "Molecular phylogeny of the lemur family Cheirogaleidae (primates)
RT based on mitochondrial DNA sequences.";
RL Mol. Phylogenet. Evol. 19:45-56(2001).
DR EMBL; AF224641; AAK70615.1; -
DR EMBL; AF224642; AAK70619.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
XW Mitochondrion.
FT NON TER
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 25.4%; Score 16; DB 8; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 2 HWSG 5
Db 5 YWVG 8

RESULT 15
Q94XE6 PRELIMINARY; PRT; 9 AA.
AC Q94XE6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome c oxidase subunit III (Fragment).
GN COX3.
OS Tectocoris diopthalmus (cotton harlequin bug).
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC Pentheteroptera; Pentatomomorpha; Pentatomidae; Pentatomidae;
OC Tectocoris.
OX NCBI_TaxID=159956;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21396409; PubMed=11504862;
RA Shao R., Campbell N.J.H., Schmidt E.R., Barker S.C.;
RT "Increased rate of gene rearrangement in the mitochondrial genomes of
RT three orders of hemipteroid insects.";
RL Mol. Biol. Evol. 18:1828-1832(2001).
DR EMBL; AF335990; AAKS283.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
XW Mitochondrion.
FT NON TER
SQ SEQUENCE 9 AA; 1206 MW; A2C563636B5041A6 CRC64;

Query Match 25.4%; Score 16; DB 8; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 2 HWSG 5
Db 5 YWVG 8

Search completed: June 1, 2004, 10:28:14
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 10:25:45 ; Search time 54 Seconds
(without alignments)
47.091 Million cell updates/sec

Title: US-09-756-899A-1

Perfect score: 63

Sequence: 1 AARSHCHCL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 231240

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	100.0	9	AAV77664	AAV77664 Immunoglobulin
2	63	100.0	9	ADC42292	ADC42292 Immunoglobulin
3	36	57.1	8	AAV21052	AAV21052 Human Gli
4	35	55.6	8	AAV49950	AAV49950 Protein d
5	35	55.6	9	AAV49952	AAV49952 Protein d
6	33	52.4	6	AAV49951	AAV49951 Protein d
7	33	52.4	6	ABP54937	ABP54937 Active si
8	33	52.4	7	AAW56315	AAW56315 Active ce
9	32	50.8	7	AAV43470	AAV43470 Flavoprot
10	32	50.8	9	AAV19859	AAV19859 TRIP zinc
11	31	49.2	9	AAW23018	AAW23018 HIV Pepti
12	31	49.2	9	AAU23853	AAU23853 Human MHC
13	31	49.2	9	AAU24423	AAU24423 Human MHC
14	30	47.6	9	AAU24453	AAU24453 Human MHC
15	30	47.6	9	ABW05276	ABW05276 Vascular
16	29	46.0	9	AAW6789	AAW6789 Lamprey I
17	29	46.0	9	AAW6787	AAW6787 Lamprey I
18	29	46.0	9	AAW50022	AAW50022 Mutant an
19	28	44.4	8	AAW85445	AAW85445 Chymotryp
20	28	44.4	9	AAW89336	AAW89336 Luteinisi
21	28	44.4	9	AAW43842	AAW43842 Specific
22	28	44.4	9	AAW22404	AAW22404 HIV pepti
23	28	44.4	9	AAW23297	AAW23297 HIV pepti
24	28	44.4	9	AAW23021	AAW23021 HIV pepti
25	28	44.4	9	AAW23295	AAW23295 HIV pepti

Abv75265 Biologica
Abv56215 Luteinisi
Add70027 Bombesin/
Aap71501 Growth ho
Aar73339 Human TSH
Aar73340 Human TSH
Aaw30454 Somatosta
Aap71502 Growth ho
Aap82065 Polypepti
Aar43763 MHC Class
Aaw3681 HPV16 E6
Aaw30461 Somatosta
Aaw54748 Peptide f
Aau02357 HLA bindi
Abj04593 Bone marr
Abj60234 184P182-r
Aay78293 Metalloth
Aab3980 Metal bin
Aab72710 Repetitiv
Abg69284 Metalloth

ALIGNMENTS

RESULT 1

AAV77664

ID AAV77664 standard; peptide; 9 AA.

XX AAV77664;

DT 12-MAY-2000 (first entry)

DE Immunoglobulin (Ig) free light chain binding peptide LCBP.

XX Immunoglobulin; Ig; light chain; LC; multiple sclerosis; antiasthmatic;
XX anti-allergy; chronic inflammatory bowel disorder; antiinflammatory;
XX antiviral; virucide; neuroprotective; antimigraine.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetylation"

FT Modified-site 9 /note= "C-terminal amide"

FT WO200002915-A1.

XX 20-JAN-2000.

XX 07-JUL-1999; 99WO-NL000430.

XX 09-JUL-1998; 98NL-01009601.

XX (UYUT-) RIJXSUNIV UTRECHT.

XX Redegeld FAM, Kraneveld AD, Nijkamp FP;

XX WPI; 2000-171132/15.

XX New compounds which inhibit immunoglobulin light chain binding to mast cells, used for treating, e.g. asthma.
XX Claim 2; Page 13; 26pp; English.

XX The invention relates to compounds which inhibit the binding of the free light chain (LC) of immunoglobulin (Ig) to mast cells. The compounds can be used for treating e.g. asthma, allergy, chronic inflammatory bowel disorders, viral infection or multiple sclerosis, and possibly migraine.
XX The methods can also be used for diagnosing a disease in a patient having an elevated level of the free LC of Ig in a body fluid. The present sequence represents a peptide capable of binding to the free LC of Ig

XX Sequence 9 AA;
SQ

Query Match 100.0%; Score 63; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AHWSGHCL 9
DB 1 AHWSGHCL 9

RESULT 2
ADC42292
ID ADC42292 standard; peptide; 9 AA.
XX
AC ADC42292;
XX
DT 18-DEC-2003 (first entry)
XX
DE Immunoglobulin light chain blocking peptide.
XX
KW gene therapy; immunoglobulin-free-light-chain receptor; Ig-LC receptor;
KW signal transduction pathway activation; gamma-chain associated receptor;
KW dermatitis; asthma; psoriasis; inflammatory bowel disease;
KW rheumatoid arthritis; Sjogren lupus erythematosus;
KW systemic lupus erythematosus; multiple sclerosis;
KW Immunoglobulin light chain.
XX
OS Unidentified.
XX
PN WO2003074563-A2.
XX
PD 12-SEP-2003.
XX
PF 05-MAR-2003; 2003WO-NL000167.
XX
PR 06-MAR-2002; 2002EP-00075874.
FR 06-MAR-2002; 2002US-0362040P.
PR 14-JUN-2002; 2002EP-00077352.
XX
PA (FORN-) FORNIX BIOSCIENCES NV.
XX
PI Nijkamp EP, Redegeld FAM, Kraneveld AD, Van De Winkel JGJ;
PI Vidarsson G;
XX
DR WPI; 2003-712878/67.
XX

New cell comprising an immunoglobulin-free-light-chain (Ig-LC) receptor that activates a signal transduction pathway in the cell, useful for selecting a compound capable of preventing binding of Ig-LC to the receptor.

PS Example 3; Page 13; 69pp; English.
PS

The invention comprises a cell containing an immunoglobulin-free-light-chain (Ig-LC) receptor that is capable of activating a signal transduction pathway in the cell upon binding of an Ig-LC to the receptor. The signal transduction is independent of the presence of a functional common gamma-chain associated receptor on the cell. The cell of the invention is useful for selecting a compound capable of preventing a binding of Ig-LC to the receptor. The cell is also useful for preparing a medicament for an animal suffering from or at risk of suffering from: dermatitis, asthma, psoriasis, inflammatory bowel disease, rheumatoid arthritis, Sjogren and systemic lupus erythematosus, and multiple sclerosis. The present amino acid sequence represents a peptide that binds to immunoglobulin light chains and therefore prevents the binding of the light chain to its receptor.

XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 63; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AHWSGHCL 9
DB 1 AHWSGHCL 9

RESULT 3
AAY21052
ID AAY21052 standard; protein; 8 AA.
XX
AC AAY21052;
XX
DT 22-JUL-1999 (first entry)
XX
DE Human glial fibrillary acidic protein GFAP mutant fragment 61.
XX
KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer; frameshift mutation; age-related disease; neurodegenerative disorder; Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal; Huntington's disease; multiple sclerosis; alcoholic liver disease; diabetes mellitus type II; microtubule associated protein; Tau; Big Tau; ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M; neurofilament-F; presenilin 1; presenilin II; cellular tumour antigen; Glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPP-1; bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A; high mobility group protein-C; neuroendocrine specific protein A.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9845322-A2.
XX
PD 15-OCT-1998.
XX
PF 02-APR-1998; 98WO-IB000705.
XX
PR 10-APR-1997; 97US-0043163P.
XX
PA (ROYA-) ROYAL NETHERLANDS ACADEMY OF ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
PA (UYUT-) RIJKSUNIV UTRECHT.
XX
PI Van Leeuwen FW, Grosveld FG, Burbach JPH;
XX
DR WPI; 1998-609901/51.
DR N-PSDB; AAX75764.
XX

Diagnosing disease by detecting frameshift mutations in RNA or corresponding protein mutations - used to diagnose cancer and neurological diseases, particularly Alzheimer's disease, and also for treatment and prevention with specific ribozymes or wild-type RNA.

PS Disclosure; Fig 13; 258pp; English.
PS

This invention describes a novel method for the diagnosis of a disease caused by, or associated with, an RNA molecule that has a frameshift mutation. The method is used to diagnose age-related diseases, especially cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's disease, multiple sclerosis, alcoholic liver disease, diabetes mellitus type II and many others listed) or susceptibility to these disorders. The method allows a definitive diagnosis of Alzheimer's disease in living patients, at an early stage. It is based on the observation that disease may be caused by mutations in RNA rather than DNA. The invention describes the use of neuronal system RNA molecules, specifically proteins including beta-amyloid precursor protein (beta-APP), the microtubule associated proteins tau and big tau, ubiquitin B, apolipoprotein E, microtubule associated protein 2 (MAP2), neurofilament-L, neurofilament-M, neurofilament-F, presenilin 1, presenilin II, glial fibrillary acidic protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma 2 (bcl-2) proto-oncogene, semaphorin III, HUPP-1, high mobility group protein-C (HMGP-C) and neuroendocrine specific protein A


```
XX Sequence 8 AA;
SQ Query Match 57.1%; Score 36; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SGHCC 8
Db 2 SGHCC 6

RESULT 4
AAB49950
ID AAB49950 standard; peptide; 8 AA.
AC AAB49950;
DT 07-MAR-2001 (first entry)
XX Protein disulfide isomerase consensus N-terminus.
DE Disulphide bond; allergy; protein disulfide isomerase variant; detergent;
XX food additive; cosmetic.
XX Mammalia.
XX WO200070064-A1.
XX 23-NOV-2000.
XX 17-MAY-2000; 2000WO-DK000265.
XX 17-MAY-1999; 99DK-00000683.
XX 18-MAY-1999; 99DK-00000689.
XX 02-JUN-1999; 99US-0137068P.
XX (NOVO ) NOVO NORDISK AS.
XX Hjordt CM;
XX WPI; 2001-070776/08.
XX Protein disulfide isomerase variant having increased reducing properties
XX and decreased redox potential than native proteins, used to reduce
XX allergenicity of allergic proteins in feed, food or cosmetic products.
XX Example 1; Page 33; 82pp; English.
XX The present invention provides variants of the Aspergillus oryzae protein
XX disulfide isomerase enzyme. These are capable of reducing disulphide
XX bonds in proteins, which may be the cause of allergies in humans. The
XX proteins can be used to reduce the allergenicity of foods, to treat
XX scleroproteins, in detergents, in food additives and in cosmetics
XX Sequence 9 AA;
Query Match 55.6%; Score 35; DB 4; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AHWGHC 7
Db 2 APWCGHC 8

RESULT 6
AAB49951
ID AAB49951 standard; peptide; 6 AA.
XX AAB49951;
XX 07-MAR-2001 (first entry)
XX Protein disulfide isomerase consensus N-terminus.
XX Disulphide bond; allergy; protein disulfide isomerase variant; detergent;
XX food additive; cosmetic.
XX Mammalia.
XX WO200070064-A1.
XX 23-NOV-2000.
XX 17-MAY-2000; 2000WO-DK000265.
XX 17-MAY-1999; 99DK-00000683.
XX 18-MAY-1999; 99DK-00000689.
XX 02-JUN-1999; 99US-0137068P.
XX (NOVO ) NOVO NORDISK AS.

XX Sequence 8 AA;
Query Match 55.6%; Score 35; DB 4; Length 8;
Best Local Similarity 71.4%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AHWGHC 7
Db 1 APWCGHC 7

RESULT 5
AAB49952
ID AAB49952 standard; peptide; 9 AA.
XX AAB49952;
XX 07-MAR-2001 (first entry)
```

XX Hjort CM;
 XX WPI; 2001-070776/08.
 XX Protein disulfide isomerase variant having increased reducing properties
 PT and decreased redox potential than native proteins, used to reduce
 PT allergenicity of allergic proteins in feed, food or cosmetic products.
 XX Example 1; Page 33; 82pp; English.
 XX The present invention provides variants of the Aspergillus oryzae protein
 CC disulfide isomerase enzyme. These are capable of reducing disulphide
 CC bonds in proteins, which may be the cause of allergies in humans. The
 CC proteins can be used to reduce the allergenicity of foods, to treat
 CC scleroproteins, in detergents, in food additives and in cosmetics
 XX Sequence 6 AA;
 SQ

Query Match 52.4%; Score 33; DB 4; Length 6;
 Best Local Similarity 80.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 WSGHC 7
 DB 1 WCGHC 5

RESULT 7
 AWP54937
 ID AWP54937 standard; peptide; 6 AA.
 XX AC AWP54937;
 XX DT 08-JAN-2003 (first entry)
 XX DE Active site peptide of protein disulfide isomerase.
 XX KW Organoarsenical; arsenoxide; antiarthritic; antitumor; osteopathic;
 XX KW antirheumatic; arthritis; therapy; protein disulfide isomerase; enzyme.
 XX OS Homo sapiens.
 XX WO200274305-A1.
 XX 26-SEP-2002.
 XX 19-MAR-2002; 2002WO-AU0000310.
 XX 19-MAR-2001; 2001AU-00003798.
 XX (UNIX) UNISEARCH LTD.
 XX Hogg PJ, Donoghue N;
 XX WPI; 2002-750519/81.
 XX Use of an impermeable cell membrane compound for the treatment of
 PT arthritis in vertebrates.
 XX Example 2(b); Page 42; 91pp; English.
 XX The present sequence is that of a hexapeptide comprising the active site
 CC sequence of protein disulfide isomerase. The hexapeptide was used in an
 CC example from the invention that examined the binding of 4-(N-(S-
 CC glutathionylacetyl)amino)phenylarsenoxide (GSAO) to peptide and protein
 CC dithiols. The results indicated that GSAO selectively binds proteins
 CC containing closely spaced thiols. The invention provides a method for the
 CC treatment and/or prophylaxis of arthritis using a compound of formula A-
 CC (L-Y)p, where A is a cell-membrane impermeable pendant group, L is a
 CC linker and/or spacer, Y is an arsenoxide or arsenoxide equivalent, and p
 CC is an integer from 1 to 10. The compound, e.g. GSAO, is useful for the
 CC treatment of arthritic conditions such as calcific periarthritis,

enteropathic, chronic, thumb, hip, peripheral, psoriatic, rheumatoid,
 CC septic or Jaccoud's arthritis, gout, hand or knee osteoarthritis,
 CC juvenile osteoarthritis, oligoarthritis, and polyarthritis (all claimed).
 CC The compound blocks angiogenesis in the synovial tissue and leucocyte
 CC ingress that triggers inflammation, with no signs or symptoms of toxicity
 XX Sequence 6 AA;
 SQ

Query Match 52.4%; Score 33; DB 5; Length 6;
 Best Local Similarity 80.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 WSGHC 7
 DB 1 WCGHC 5

RESULT 8
 AAW56315
 ID AAW56315 standard; protein; 7 AA.
 XX AC AAW56315;
 XX DT 24-JUL-1998 (first entry)
 XX DE Active centre of the disulphide bond exchange reaction catalysed by PDI.
 XX KW Protein disulphide isomerase; PDI; methylotrophic yeast;
 XX KW Candida boidinii strain S2; endoplasmic reticulum; stable conformation;
 XX KW recombination; disulphide bond; secretory protein; gene therapy;
 XX KW endoplasmic reticulum retention signal sequence; genetic engineering.
 XX OS Candida boidinii.
 XX EP828004-A2.
 XX 11-MAR-1998.
 XX 04-SEP-1997; 97EP-00306871.
 XX 04-SEP-1996; 96JP-00234287.
 XX (SUNR) SUNTORY LTD.
 XX Sakai Y, Kato N, Shibano Y;
 XX WPI; 1998-161102/15.
 XX Methylotrophic yeast protein disulphide isomerase - and corresponding
 PT gene useful for increasing yields of secreted heterologous proteins in
 PT Candida boidinii.
 XX Example 1; Page 4; 30pp; English.
 XX AAW56315-16 represent the active centre of the exchange reaction of
 CC disulphide bonds, and are found in the protein disulphide isomerase (PDI)
 CC derived from Candida boidinii. The sequences are related to AAW56314. PDI
 CC is a major protein present in the lumen of the endoplasmic reticulum. PDI
 CC is believed to be an enzyme which catalyses formation of stable
 CC conformation by recombining disulphide bonds of secretory proteins. As
 CC the protein must stay in the endoplasmic reticulum, it has a sequence
 CC known as the endoplasmic reticulum retention signal sequence at the
 CC carboxy terminus (AAW56317). The PDI protein is useful in gene therapy
 CC and genetic engineering. The PDI gene may be coexpressed with a gene of
 CC interest to ensure the production of a correctly folded biologically
 CC active protein
 XX Sequence 7 AA;
 SQ

Query Match 52.4%; Score 33; DB 2; Length 7;
 Best Local Similarity 80.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      3 WSGHC 7
DB      2 WCGHC 6

RESULT 9
ID AAY43470 standard; peptide; 7 AA.
XX
AC AAY43470;
XX
XX 27-AUG-2003 (revised)
DT 17-DEC-1999 (first entry)
XX
DE Flavoprotein EpiD substrate determination peptide #31.
XX
KW Staphylococcus epidermis; epidermin; epiA; epiB; epiC; epiD; epiP; epiQ;
KW epiY; epiZ; oxidative carboxylation; flavoprotein.
XX
OS Synthetic.
OS Staphylococcus epidermidis.
XX
XX US9562253-A.
XX
XX 05-OCT-1999.
XX
XX 13-MAY-1996; 96US-00645193.
XX
XX 13-MAY-1996; 96US-00645193.
XX
XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
XX Kupke T, Kempter C, Jung G, Goetz F;
XX
XX WPI; 1999-589714/50.
XX
XX Oxidative decarboxylation of peptides by the flavoprotein EpiD.
XX
XX Example 11; Col 37; 89pp; English.
XX
CC The present invention describes the oxidative decarboxylation of a
CC peptide comprising combining the peptide which contains at its carboxy
CC terminus the amino acid sequence (I) with EpiD; X1-X2-X3-X4-X5-X6-C (I)
CC where X1, X2, X3 and X4 = any one of the 20 common amino acids; X5 = Tyr,
CC Val, Met, Phe, Ile, Leu or Trp; and X6 = Cys, Ala, Ser, Val or Thr
CC providing that the carboxy terminus of the peptide is not sequence (2)-
CC (5): SFNSVCC (2); SFNSFCC (3); SFNSWCC (4) and SFNSVSC (5). The method
CC can be used for the preparation of derivatives of known peptides and
CC hormones with a cysteine residue in the unmodified peptide replaced by
CC sulphide-bridged amino acids and serine and thiamine replaced by
CC corresponding dehydroamino acid residues. The formation of novel
CC compounds for experimental purposes or for the formation of known
CC compounds or their derivatives in new hosts e.g. biologically active
CC peptide derivatives containing dehydroamino acid residues and/or at least
CC one lanthionine bridge and/or methyl-lanthionine bridge such as
CC derivatives of human insulin, oxytocin, vasopressin, peptide antibiotics,
CC hormone inhibitors such as elastase inhibitor and fibrinolytically
CC active agents such as human tissue plasminogen activator can also be
CC achieved using this method. Derivatives modified by this method retain
CC the biological activity of the parent compound but have increased
CC stability and improved half-lives. AAY43470 to AAY43475, and AAY25496 to
CC AA225506, represent sequences used in the exemplification of the present
CC invention. (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 7 AA;
XX
Query Match 50.8%; Score 32; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 WSGHC 8
DB      2 WNSVCC 7

RESULT 10
ID AAE19859 standard; peptide; 9 AA.
XX
AC AAE19859;
XX
XX 18-JUN-2002 (first entry)
DT
DE TRIP zinc finger domain consensus.
XX
XX Human; tumour necrosis factor; TNF; TNF receptor associated factor; TRAF;
KW TRAF interacting protein; TRIP; cell activation; cell proliferation;
KW cell death; therapy; cytostatic; zinc finger domain.
XX
XX Homo sapiens.
OS Mus sp.
XX
XX US6346605-B1.
XX
XX 12-FEB-2002.
XX
XX 31-MAR-1998; 98US-00052089.
XX
XX 01-APR-1997; 97US-0042293P.
XX
XX 07-APR-1997; 97US-0042747P.
XX
XX (UYRQ ) UNIV ROCKEFELLER.
XX
XX Lee SY, Choi Y;
XX
XX WPI; 2002-225005/28.
XX
XX New tumor necrosis factor receptor associated factor interacting protein.
XX useful for inhibiting NF-kappa B activation, and for modulating signals
XX responsible for cell activation, cell proliferation and cell death.
XX
XX Example 2; Fig 2D; 37pp; English.
XX
CC The present invention relates to a tumour necrosis factor (TNF) receptor
CC associated factor (TRAF) interacting protein (TRIP), which is a regulator
CC capable of binding to TRAF2. TRIP is useful for inhibiting NF-kappa B
CC activation and for modulating signals responsible for cell activation,
CC cell proliferation and cell death. Thus, TRIP is useful for treating
CC diseases associated with altered cell proliferation and cell death. The
CC present sequence is zinc finger domain consensus sequence of human and
CC mouse TRIP
XX
XX Sequence 9 AA;
XX
Query Match 50.8%; Score 32; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 GHCC 8
DB      4 GHCC 7

RESULT 11
ID AAM23018 standard; peptide; 9 AA.
XX
AC AAM23018;
XX
XX 22-OCT-2001 (first entry)
DT
DE HIV peptide SEQ ID NO 903.
XX
XX Cytotoxic T cell lymphocyte; CTL; anti-HIV; viral disease; HIV;
KW human immunodeficiency virus; vaccine.
XX

```

OS Human immunodeficiency virus.
 OS Synthetic.
 PN WO200155177-A2.
 XX
 XX 02-AUG-2001.
 XX
 XX 29-JAN-2001; 2001WO-DK000059.
 XX
 XX 28-JAN-2000; 2000EP-00610017.
 PR 31-JAN-2000; 2000US-0179333P.
 XX
 XX (STAT-) SPATENS SERUM INST.
 XX
 XX Ponsgaard A, Brunak S, Buus S, Corbet S, Lauemoller SL, Hansen J;
 PI WPI; 2001-476184/51.
 XX
 XX The generation of cytotoxic T cell lymphocytes epitopes for use in anti-
 XX HIV vaccines.
 PT
 XX
 XX Example 4; Page 70; 383pp; English.
 XX
 XX The invention relates to identification of cytotoxic T cell lymphocyte
 CC (CTL) epitopes (AAM22116-AAM23484) that generate anti-HIV activity. CTL
 CC are a major protective mechanism against viral diseases. Antibodies may
 CC neutralise extracellular human immunodeficiency virus (HIV) and limit or
 CC prevent infection of cells in the host, but CTL will limit viral
 CC production by killing the cell. The CTL epitopes are useful in medicine,
 CC in the manufacture of vaccines or diagnostic agents
 CC
 XX Sequence 9 AA;
 SQ

Query Match 49.2%; Score 31; DB 4; Length 9;
 Best Local Similarity 57.1%; Pred. No. 1.4e+06;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 WSGHCCL 9
 | : | | |
 DQ 3 WATHACL 9

RESULT 12
 AAU23853
 ID AAU23853 standard; peptide; 9 AA.
 XX
 XX AAU23853;
 AC
 XX 17-DEC-2001 (first entry)
 DT
 XX Human MHC class I molecule HLA-A1 binding 103P2D6 peptide #38.
 DE
 XX

103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;
 KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;
 KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;
 KW single chain monoclonal antibody; cervix; human.
 XX
 XX Homo sapiens.
 OS
 XX WO200162925-A2.
 PN
 XX 30-AUG-2001.
 PD
 XX 26-FEB-2001; 2001WO-US005996.
 PF
 XX 24-FEB-2000; 2000US-0184558P.
 PR 13-JUL-2000; 2000US-0218856P.
 PR
 XX (UROC-) UROGENESYS INC.
 PA
 XX Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;
 PI Challita-Eid PM, Faris M, Jakobovits A;
 XX

DR WPI; 2001-557705/62.
 XX
 XX New polynucleotide for treating and diagnosing prostate cancer is the
 PT 103P2D6 gene which encodes for 103P2D6-related proteins.
 PT
 XX
 XX Example 15; Page 81; 132pp; English.
 XX
 XX Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and
 CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal
 CC adult tissue but is aberrantly expressed in some foetal tissues and many
 CC cancers including tumours of the prostate, testis, bladder, bone, cervix,
 CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its
 CC related protein and also peptide fragments of the protein are therefore
 CC useful for diagnosing and treating cancer. A vector comprising a
 CC polynucleotide which encodes a single chain monoclonal antibody, that
 CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme
 CC are both useful in the preparation of a composition for treating a
 CC patient with a cancer that expresses 103P2D6. The sequences can be used
 CC in diagnostic methods to monitor the level of 103P2D6 gene products in
 CC serum, blood, urine and tissue and to thereby detect the presence of
 CC cancerous cells
 CC
 XX Sequence 9 AA;
 SQ

Query Match 49.2%; Score 31; DB 4; Length 9;
 Best Local Similarity 71.4%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WSGHCCL 9
 | : | | |
 DQ 1 WSGRCGL 7

RESULT 13
 AAU24423
 ID AAU24423 standard; peptide; 9 AA.
 XX
 XX AAU24423;
 AC
 XX 17-DEC-2001 (first entry)
 DT
 XX Human MHC molecule HLA-B35 binding 103P2D6 peptide #8.
 DE
 XX

103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;
 KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;
 KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;
 KW single chain monoclonal antibody; cervix; human.
 XX
 XX Homo sapiens.
 OS
 XX WO200162925-A2.
 PN
 XX 30-AUG-2001.
 PD
 XX 26-FEB-2001; 2001WO-US005996.
 PF
 XX 24-FEB-2000; 2000US-0184558P.
 PR 13-JUL-2000; 2000US-0218856P.
 PR
 XX (UROC-) UROGENESYS INC.
 PA
 XX Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;
 PI Challita-Eid PM, Faris M, Jakobovits A;
 XX

WPI; 2001-557705/62.
 XX
 XX New polynucleotide for treating and diagnosing prostate cancer is the
 PT 103P2D6 gene which encodes for 103P2D6-related proteins.
 PT
 XX
 XX Example 15; Page 97; 132pp; English.
 XX
 XX Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and

CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal
 CC adult tissue but is aberrantly expressed in some foetal tissues and many
 CC cancers including tumours of the prostate, testis, bladder, bone, cervix,
 CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its
 CC related protein and also peptide fragments of the protein are therefore
 CC useful for diagnosing and treating cancer. A vector comprising a
 CC polynucleotide which encodes a single chain monoclonal antibody, that
 CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme
 CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,
 CC are both useful in the preparation of a composition for treating a
 CC patient with a cancer that expresses 103P2D6. The sequences can be used
 CC in diagnostic methods to monitor the level of 103P2D6 gene products in
 CC serum, blood, urine and tissue and to thereby detect the presence of
 CC cancerous cells

XX SQ Sequence 9 AA;

Query Match 49.2%; Score 31; DB 4; Length 9;
 Best Local Similarity 71.4%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WSGHCC 9
 |||||
 Db 1 WSGRCGL 7

RESULT 14
 AAU24453
 ID AAU24453 standard; peptide; 9 AA.

AC AAU24453;

XX 17-DEC-2001 (first entry)

XX Human MHC molecule HLA-B35 binding 103P2D6 peptide #38.

XX 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;
 KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;
 KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;
 KW single chain monoclonal antibody; cervix; human.

XX Homo sapiens.

OS WO200162925-A2.

PN 30-AUG-2001.

XX 26-FEB-2001; 2001WO-US005996.

XX 24-FEB-2000; 2000US-0184558P.

PR 13-JUL-2000; 2000US-0218956P.

XX (UROC-) UROGENESYS INC.

XX Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;
 PI Chailita-Eid PM, Paris M, Jakobovits A;

XX WPI; 2001-557705/62.

XX New polynucleotide for treating and diagnosing prostate cancer is the
 PT 103P2D6 gene which encodes for 103P2D6-related proteins.

XX Example 15; Page 98; 132pp; English.

XX Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and
 CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal
 CC adult tissue but is aberrantly expressed in some foetal tissues and many
 CC cancers including tumours of the prostate, testis, bladder, bone, cervix,
 CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its
 CC related protein and also peptide fragments of the protein are therefore
 CC useful for diagnosing and treating cancer. A vector comprising a
 CC polynucleotide which encodes a single chain monoclonal antibody, that
 CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme

CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,
 CC are both useful in the preparation of a composition for treating a
 CC patient with a cancer that expresses 103P2D6. The sequences can be used
 CC in diagnostic methods to monitor the level of 103P2D6 gene products in
 CC serum, blood, urine and tissue and to thereby detect the presence of
 CC cancerous cells

XX SQ Sequence 9 AA;

Query Match 47.6%; Score 30; DB 4; Length 9;
 Best Local Similarity 80.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WSGHC 7
 |||||
 Db 5 WSGRC 9

RESULT 15

ABB05276

ID ABB05276 standard; peptide; 9 AA.

XX ABB05276;

XX 29-AUG-2003 (revised)

DT 04-APR-2003 (first entry)

XX Vascular endothelial growth factor binding peptide BP 86 SEQ ID NO:108.

XX Human; tumour necrosis factor alpha; TNF-alpha; VEGF; detergent; stain;
 KW bacteriophage; phage library; vascular endothelial growth factor;
 KW collar soil; polyurethane; egg; tea; hair; skin; cleaning composition.

XX Unidentified bacteriophage.

OS Unidentified.

XX WO200179479-A2.

XX 25-OCT-2001.

PD 11-APR-2001; 2001WO-US011811.

PF 14-APR-2000; 2000US-0197259P.

XX (GENV) GENENCOR INT INC.

XX Estell DA, Murray CJ, Tijerina P, Chen Y;

XX WPI; 2002-139323/18.

XX Screening ligand library comprises allowing binding of ligand with anti-
 CC target, contacting unbound ligands with selected target to form target-
 PT bound ligand complex and identifying target bound ligands on the complex.

XX Claim 23; Page 28; 51pp; English.

XX The present invention describes a method for screening a ligand library
 CC (LL). The method comprises: (a) contacting the LL with an anti-target
 CC (AT) to allow the ligands to bind to the AT; (b) separating unbound
 CC ligands; (c) contacting the unbound ligands with a selected target (T) to
 CC allow binding of unbound ligands to (T) to form a (T)-bound ligand
 CC complex (C); (d) separating (C) from ligands which do not bind (T); and
 CC (e) identifying (T)-bound ligands on (C). The method can be used for
 CC screening a ligand library, e.g., a library of peptides, polypeptides,
 CC non-polypeptides or oligonucleotides. A ligand (L) identified by the
 CC method can be used in a cleaning, therapeutic or personal care
 CC application. The method is preferably useful for identifying peptide
 CC useful in cleaning compositions, which involves contacting peptide
 CC library with AT such as fabric, ceramic, glass, stainless steel or
 CC plastic; separating unbound AT peptides; contacting unbound AT peptides
 CC with a target which is a stain such as porphyrin derived stain; tannin
 CC derived stain, carotenoid pigment derived stain, anthocyanin pigment
 CC derived stain, soil-based derived stain, oil-based derived stain, and

CC human body stain, to allow unbound peptide to bind with the stain to form
CC stain-bound peptide complex and identifying the stain-bound peptide on
CC the stain-bound peptide complex. A selective targeting method for
CC screening a library of ligands that bind to a target may be used to
CC identify ligands that bind to a target under harsh conditions. The
CC selective targeting method may be used to screen and identify a ligand
CC useful for therapeutic intervention, e.g., a library of ligands may be
CC screened to identify a tumour-bound ligand. The selective targeting
CC method may be used to identify cell type specific surface molecules.
CC Preferred anti-targets include one or more different cell types, cells in
CC different states, or cells that do not display the surface molecule.
CC ABB05232 to ABB05346 represent phage-peptides ligands which are used in
CC the exemplification of the present invention. (Updated on 29-AUG-2003 to
CC standardise OS field)

XX
SQ Sequence 9 AA;

Query Match 47.6%; Score 30; DB 5; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AHWSGHCC 8
| | | |
Db 2 APWNSHC 9

Search completed: June 1, 2004, 10:26:56
Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 1, 2004, 10:28:21 ; Search time 42 Seconds
(without alignments)
59,833 Million cell updates/sec

Title: US-09-756-899A-1
Sequence: 1 AHWSGHCHCL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues
Total number of hits satisfying chosen parameters: 110747

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/FCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	63	100.0	9	9	US-09-756-899A-1
2	36	57.1	5	12	US-10-436-549-494
3	31	49.2	9	10	US-09-793-451-52
4	31	49.2	9	10	US-09-793-451-622
5	31	49.2	9	12	US-10-182-252A-903
6	31	49.2	9	14	US-10-283-722-52
7	31	49.2	9	15	US-10-283-722-622
8	31	49.2	9	15	US-10-283-903-52
9	31	49.2	9	15	US-10-283-903-622
10	30	47.6	5	12	US-10-436-549-497
11	30	47.6	9	9	US-09-832-723-108
12	30	47.6	9	10	US-09-793-451-652
13	30	47.6	9	14	US-10-303-331-108
14	30	47.6	9	14	US-10-283-722-652
15	30	47.6	9	15	US-10-283-903-652

Sequence 46, Appl
Sequence 289, App
Sequence 906, App
Sequence 1180, Ap
Sequence 1182, Ap
Sequence 118, App
Sequence 117, Appl
Sequence 87, Appl
Sequence 107, App
Sequence 104, App
Sequence 70, Appl
Sequence 23, Appl
Sequence 109, App
Sequence 6, Appl
Sequence 2, Appl
Sequence 6, Appl
Sequence 116, App
Sequence 217, App
Sequence 226, App
Sequence 306, App
Sequence 331, App
Sequence 413, App
Sequence 502, App
Sequence 531, App
Sequence 610, App
Sequence 63, Appl
Sequence 89, Appl
Sequence 151, App

US-08-821-739A-46
US-10-182-252A-289
US-10-182-252A-906
US-10-182-252A-1180
US-10-182-252A-1182
US-10-254-446A-118
US-10-360-101-117
US-09-911-838-87
US-10-096-986-107
US-10-396-122-104
US-10-430-685-70
US-10-197-927-23
US-09-832-723-109
US-10-216-810A-6
US-10-303-331-109
US-10-362-768A-2
US-09-938-114-6
US-09-942-052-116
US-09-942-052-217
US-09-942-052-226
US-09-942-052-306
US-09-942-052-331
US-09-942-052-413
US-09-942-052-502
US-09-942-052-531
US-09-942-052-610
US-10-357-173-63
US-10-455-720-63
US-10-360-101-89
US-10-254-446A-151

ALIGNMENTS

RESULT 1
US-09-756-899A-1
; Sequence 1, Application US/09756899A
; Patent No. US20020045186A1
; GENERAL INFORMATION:
; APPLICANT: Redgelo, Francisus
; APPLICANT: Kraneveld, Aletta
; APPLICANT: Nijkamp, Francisus
; TITLE OF INVENTION: INHIBITION OF PROTEIN BINDING TO MAST CELLS
; FILE REFERENCE: 2183-4692
; CURRENT APPLICATION NUMBER: US/09/756,899A
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: PCT/NL99/00430
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: peptide
US-09-756-899A-1

Query Match 100.0%; Score 63; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No 1e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AHWSGHCHCL 9
|||||||
DB 1 AHWSGHCHCL 9
RESULT 2
US-10-436-549-494
; Sequence 494, Application US/10436549
; Publication No. US20040038307A1
; GENERAL INFORMATION:

```
; APPLICANT: Lee, Frank D.
; APPLICANT: Meng, Dr. Xun
; APPLICANT: Chan, John W.
; APPLICANT: Zhang, Shengsheng J.
; APPLICANT: Benkovic, Stephen J.
; TITLE OF INVENTION: UNIQUE RECOGNITION SEQUENCES AND METHODS OF USE THEREOF IN
; FILE REFERENCE: ENGE-P01-001
; CURRENT APPLICATION NUMBER: US/10/436,549
; CURRENT FILING DATE: 2003-05-12
; PRIOR APPLICATION NUMBER: 60/379,626
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/393,233
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,235
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,211
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,280
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,197
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,223
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/430,948
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/433,319
; PRIOR FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/393,137
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 614
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 494
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Human
US-10-436-549-494

Query Match          57.1%; Score 36; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      5 GHCCCL 9
Db      1 GHCCCL 5

RESULT 3
US-09-793-451-52
; Sequence 52, Application US/09793451
; Publication No. US20030157597A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E. H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Ava Jakobovits
; TITLE OF INVENTION: 103P2D6; TISSUE SPECIFIC PROTEIN HIGHLY
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-622

Query Match          49.2%; Score 31; DB 10; Length 9;
Best Local Similarity 71.4%; Pred. No. 1e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      3 WSGHCCL 9
Db      1 WSGRCGL 7

RESULT 5
US-10-182-252A-903
; Sequence 903, Application US/10182252A
; Publication No. US20040072162A1
; GENERAL INFORMATION:
; APPLICANT: FOMSGAARD, ANDERS
; APPLICANT: BRUNAK, SOREN
; APPLICANT: BUUS, SOREN
; APPLICANT: CORBET, SYLVIE
; APPLICANT: LAUEMOLLER, SANNE LISE
; APPLICANT: HANSEN, JAN
; TITLE OF INVENTION: HIV PEPTIDE AND NUCLEIC ACIDS ENCODING THEM FOR DIAGNOSIS AND
; FILE REFERENCE: 030307/0205
; CURRENT APPLICATION NUMBER: US/10/182,252A
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: PCT/DK01/00059
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: EP 00610017.6
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/179,333
```

```
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-52

Query Match          49.2%; Score 31; DB 10; Length 9;
Best Local Similarity 71.4%; Pred. No. 1e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      3 WSGHCCL 9
Db      1 WSGRCGL 7

RESULT 4
US-09-793-451-622
; Sequence 622, Application US/09793451
; Publication No. US20030157597A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E. H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Ava Jakobovits
; TITLE OF INVENTION: 103P2D6; TISSUE SPECIFIC PROTEIN HIGHLY
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 622
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-622

Query Match          49.2%; Score 31; DB 10; Length 9;
Best Local Similarity 71.4%; Pred. No. 1e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      3 WSGHCCL 9
Db      1 WSGRCGL 7

RESULT 5
US-10-182-252A-903
; Sequence 903, Application US/10182252A
; Publication No. US20040072162A1
; GENERAL INFORMATION:
; APPLICANT: FOMSGAARD, ANDERS
; APPLICANT: BRUNAK, SOREN
; APPLICANT: BUUS, SOREN
; APPLICANT: CORBET, SYLVIE
; APPLICANT: LAUEMOLLER, SANNE LISE
; APPLICANT: HANSEN, JAN
; TITLE OF INVENTION: HIV PEPTIDE AND NUCLEIC ACIDS ENCODING THEM FOR DIAGNOSIS AND
; FILE REFERENCE: 030307/0205
; CURRENT APPLICATION NUMBER: US/10/182,252A
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: PCT/DK01/00059
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: EP 00610017.6
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/179,333
```



```
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 1388
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 903
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV peptide
US-10-182-252A-903
```

```
Query Match          49.2%; Score 31; DB 12; Length 9;
Best Local Similarity 57.1%; Pred. No. 1e+06; 2; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 WSGHCCL 9
   | | | |
DB 3 WATHACL 9
```

```
RESULT 6
US-10-283-722-52
; Sequence 52, Application US/10283722
; Publication No. US20030194407A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/10/283,722
; PRIOR FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/793,451
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-283-722-52
```

```
Query Match          49.2%; Score 31; DB 14; Length 9;
Best Local Similarity 71.4%; Pred. No. 1e+06; 2; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 WSGHCCL 9
   | | | |
DB 1 WSGRCGL 7
```

```
RESULT 7
US-10-283-722-622
; Sequence 622, Application US/10283722
; Publication No. US20030194407A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Mary Faris
```

```
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/10/283,722
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/793,451
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 622
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-283-722-622
```

```
Query Match          49.2%; Score 31; DB 14; Length 9;
Best Local Similarity 71.4%; Pred. No. 1e+06; 2; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 WSGHCCL 9
   | | | |
DB 1 WSGRCGL 7
```

```
RESULT 8
US-10-283-903-52
; Sequence 52, Application US/10283903
; Publication No. US20030219766A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/10/283,903
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/793,451
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-283-903-52
```

```
Query Match          49.2%; Score 31; DB 15; Length 9;
Best Local Similarity 71.4%; Pred. No. 1e+06; 2; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 WSGHCCL 9
   | | | |
DB 1 WSGRCGL 7
```

```
RESULT 9
US-10-283-903-622
; Sequence 622, Application US/10283903
```

Publication No. US20030219766A1
GENERAL INFORMATION:
APPLICANT: Arthur B. Raitano
APPLICANT: Daniel E.H. Afar
APPLICANT: Gazelle S. Rastegar
APPLICANT: Steve Chappell Mitchell
APPLICANT: Rene S. Hubert
APPLICANT: Pia M. Challita-Eid
APPLICANT: Mary Faris
APPLICANT: Aya Jakobovits
TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
FILE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
FILE REFERENCE: 129.2USU2
CURRENT APPLICATION NUMBER: US/10/283.903
CURRENT FILING DATE: 2003-02-03
PRIOR APPLICATION NUMBER: US/09/793.451
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: 60/184,558
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/218,856
PRIOR FILING DATE: 2000-07-13
NUMBER OF SEQ ID NOS: 752
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 622
LENGTH: 9
TYPE: PRT
ORGANISM: homo sapiens
US-10-283-903-622

Query Match 49.2%; Score 31; DB 15; Length 9;
Best Local Similarity 71.4%; Pred. No. 1e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WSGHCC 9
DB 1 WSGRGL 7

RESULT 10
US-10-436-549-497
Sequence 497, Application US/10436549
Publication No. US20040038307A1
GENERAL INFORMATION:
APPLICANT: Lee, Frank D.
APPLICANT: Meng, Dr. Xun
APPLICANT: Chan, John W.
APPLICANT: Zhang, Shengsheng
APPLICANT: Benkovic, Stephen J.
TITLE OF INVENTION: UNIQUE RECOGNITION SEQUENCES AND METHODS OF USE THEREOF IN
FILE OF INVENTION: PROTEIN ANALYSIS
FILE REFERENCE: ENGE-P01-001
CURRENT APPLICATION NUMBER: US/10/436.549
CURRENT FILING DATE: 2003-05-12
PRIOR APPLICATION NUMBER: 60/379,626
PRIOR FILING DATE: 2002-05-10
PRIOR APPLICATION NUMBER: 60/393,233
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: 60/393,235
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: 60/393,211
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: 60/393,280
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: 60/393,197
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: 60/393,223
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: 60/430,948
PRIOR FILING DATE: 2002-12-04
PRIOR APPLICATION NUMBER: 60/433,319
PRIOR FILING DATE: 2002-12-13
PRIOR APPLICATION NUMBER: 60/393,137
PRIOR FILING DATE: 2002-07-01

NUMBER OF SEQ ID NOS: 614
SOFTWARE: PatentIn version 3.2
SEQ ID NO 497
LENGTH: 5
TYPE: PRT
ORGANISM: Human
US-10-436-549-497

Query Match 47.6%; Score 30; DB 12; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HCCL 9
DB 1 HCCL 4

RESULT 11
US-09-832-723-108
Sequence 108, Application US/09832723
Patent No. US20020098524A1
GENERAL INFORMATION:
APPLICANT: Estell, David A.
APPLICANT: Chen, Yiyou
APPLICANT: Murray, Christopher J.
APPLICANT: Tijerina, Pilar
TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
FILE REFERENCE: GC617-2
CURRENT APPLICATION NUMBER: US/09/832,723
CURRENT FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: US 60/197,259
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 117
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 108
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptides screened from a phage display random
OTHER INFORMATION: Peptide library
US-09-832-723-108

Query Match 47.6%; Score 30; DB 9; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AHWGHC 8
DB 2 APWNSHC 9

RESULT 12
US-09-793-451-652
Sequence 652, Application US/09793451
Publication No. US20030157597A1
GENERAL INFORMATION:
APPLICANT: Arthur B. Raitano
APPLICANT: Daniel E.H. Afar
APPLICANT: Gazelle S. Rastegar
APPLICANT: Steve Chappell Mitchell
APPLICANT: Rene S. Hubert
APPLICANT: Pia M. Challita-Eid
APPLICANT: Mary Faris
APPLICANT: Aya Jakobovits
TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
FILE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
FILE REFERENCE: 129.2USU2
CURRENT APPLICATION NUMBER: US/09/793.451
CURRENT FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: 60/184,558
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/218,856

```

; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 652
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-652

Query Match      47.6%; Score 30; DB 10; Length 9;
Best Local Similarity 80.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 WSGHC 7
DB      5 WSGRC 9

RESULT 13
US-10-303-331-108
; Sequence 108, Application US/10303331
; Publication No. US20030152976A1
; GENERAL INFORMATION:
; APPLICANT: Janssen, Giselle G.
; APPLICANT: Murray, Christopher J.
; APPLICANT: Winetzky, Deborah S.
; TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
; FILE REFERENCE: GC617-3
; CURRENT APPLICATION NUMBER: US/10/303,331
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: US 09/832,723
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,259
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptides screened from a phage display random
; OTHER INFORMATION: Peptide library
US-10-303-331-108

Query Match      47.6%; Score 30; DB 14; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+06;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 AHWSGHCC 8
DB      2 APWNSHC 9

RESULT 14
US-10-283-722-652
; Sequence 652, Application US/10283722
; Publication No. US2003019407A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/10/283,722
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/793,451
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 652
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-283-722-652

Query Match      47.6%; Score 30; DB 15; Length 9;
Best Local Similarity 80.0%; Pred. No. 1e+06;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 WSGHC 7
DB      5 WSGRC 9

Search completed: June 1, 2004, 10:33:57
Job time : 43 secs
```